

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO SÓCIO-ECONÔMICO
DEPARTAMENTO DE CIÊNCIAS DA ADMINISTRAÇÃO
COORDENADORIA DE ESTÁGIO

UM ESTUDO SOBRE *DATA WAREHOUSING*

ANA LUCIA DOS SANTOS

Florianópolis, novembro de 1999

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO SÓCIO-ECONÔMICO
DEPARTAMENTO DE CIÊNCIAS DA ADMINISTRAÇÃO
COORDENADORIA DE ESTÁGIO

UM ESTUDO SOBRE *DATA WAREHOUSING*

ANA LUCIA DOS SANTOS

PROFESSORA ORIENTADORA
ALESSANDRA DE LINHARES JACOBSEN, MSc

ÁREA DE CONCENTRAÇÃO
ADMINISTRAÇÃO E INFORMÁTICA

Florianópolis, novembro de 1999

Este Trabalho de Conclusão de Estágio foi apresentado e julgado perante a Banca Examinadora que atribui a nota à aluna Ana Lucia dos Santos na disciplina Estágio Supervisionado Obrigatório - CAD 5401

Banca Examinadora

Prof. Mário de Souza Almeida, MSc

Prof. Pedro Carlos Schenini, Dr

Profª Alessandra de Linhares Jacobsen, MSc
Presidente

À minha família.

AGRADECIMENTOS

À prof^a Alessandra de Linhares Jacobsen, pela orientação, amizade, paciência e incentivo.

Aos professores do curso, pela disposição em transmitir a própria experiência e em despertar nos seus alunos a capacidade de tratar novos problemas e desenvolverem-se sozinhos, o que têm sido um grande estímulo para nós (seus alunos).

A amiga Teresinha Batista Nunes, pelo auxílio, na realização deste trabalho, que sempre me influenciou positivamente.

A minha irmã Luciana Maria dos Santos, pelo seu apoio e incentivo, desde o vestibular até a conclusão deste trabalho.

Aos colegas do Tribunal Regional do Trabalho 12^a Região, pela disposição e colaboração.

EPÍGRAFE

“Não me venham com conclusões!

A única conclusão é morrer”

Álvaro de Campos

SUMÁRIO

RESUMO	v
1. INTRODUÇÃO	1
1.1 Apresentação do Tema-Problema.....	3
1.2 Justificativa.....	5
1.2.1 Importância.....	5
1.2.2 Oportunidade	6
1.2.3 Viabilidade	7
2. OBJETIVOS	8
2.1 Objetivo geral	8
2.2 Objetivos específicos.....	8
3. FUNDAMENTAÇÃO TEÓRICA.....	10
3.1 A crescente necessidade de planejamento e informação	10
3.2 A tecnologia de informação como apoio à decisão e distribuição de informações	15
3.3 Fontes e formas de recuperação de informações gerenciais	23
3.4 <i>Data warehousing</i>	27
3.5 <i>Data marts</i>	33
3.6 Definição de um modelo de <i>data warehouse</i>	35
3.7 Ferramentas para recuperação e análise de dados	39
3.7.1 Consultas convencionais.....	40
3.7.2 OLAP.....	40
3.11.3 <i>Data mining</i>	41
3.12 <i>Data warehouse e Web</i>	42
4. METODOLOGIA	44
4.1 Delineamento de pesquisa.....	44
4.2 Delimitação da população pesquisada	45
4.3 Técnicas de coleta de dados	45

4.4 Tratamento dos dados	45
4.5 Limitações do método	46
5. APRESENTAÇÃO E ANÁLISE DOS DADOS.....	47
5.1 Caracterização da organização	47
5.2 Definição do modelo	53
5.2.1 Definição dos processos	53
5.2.2 A granularidade das tabelas de fatos	57
5.2.3 As dimensões de cada tabela de fatos.....	58
5.2.4 Definição dos fatos armazenados	61
5.2.5 Os atributos de cada dimensão	61
5.2.6 Forma de tratamento das dimensões de modificação lenta.....	64
5.2.7 Definição de agregados	65
5.2.8 Amplitude de tempo do histórico armazenado no data mart	65
5.2.9 Intervalos de extração de dados e carga do data mart	65
5.2.10 A disponibilização dos dados	66
5.3 Limitações do projeto	67
5.4 Tendências da área no Tribunal.....	67
CONSIDERAÇÕES FINAIS	69
REFERÊNCIAS BIBLIOGRÁFICAS	71

RESUMO

*Like nature's rain, information is the
very lifeblood of the enterprise.*

www.redwoodsw.com

Este trabalho parte da importância da informação. Nas empresas, atualmente, independente da área de atuação, do nível de concorrência, de visar lucro ou não, de ser pública ou privada, a disponibilidade de informações torna-se cada vez mais crítica. São necessárias informações sobre a empresa, como ela está funcionando, custos, processos, projetos, investimentos, resultados, os concorrentes, o ambiente e os mercados. Neste contexto, tem recebido destaque dentre as tecnologias voltadas ao suprimento de informações gerenciais e o suporte à decisão, as tecnologias de *data warehousing*. *Data warehousing*, pode ser definido como a tecnologia que atualmente tem proporcionado melhorias no processo de tomada de decisões. Um *data warehouse* contém uma grande quantidade e variedade de dados capazes de reproduzirem uma imagem coerente das condições do negócio em um dado momento. Um *data warehouse* é informacional ao invés de operacional, seu objetivo é reunir dados operacionais e de fontes externas e torna-las acessíveis para entendimento, gerência e uso. Neste trabalho, procurou-se definir estas tecnologias através de uma revisão bibliográfica do assunto e contextualiza-las através da elaboração de um modelo. Para a elaboração deste modelo, foi utilizada a área de informática do Tribunal Regional do Trabalho - 12ª Região. Foi desenvolvida uma pesquisa exploratória aplicada, com abordagem qualitativa, sendo ao mesmo tempo uma pesquisa bibliográfica e um estudo de caso, ao explorar o tema e aplica-lo na produção de um modelo. O Tribunal, como empresa pública voltada ao atendimento à comunidade, tem buscado empregar as tecnologias disponíveis no mercado na agilização de suas atividades, racionalização do uso dos recursos e democratização do acesso à seus serviços e bancos de dados. A elaboração e posterior implantação do modelo proposto neste trabalho, ou seja, de um modelo de *data mart* para suprimento a necessidades de informação da área de informática do TRT vem de encontro a estas políticas.

1. INTRODUÇÃO

De modo geral, pode-se dizer que a informatização das empresas tem se apresentado como um processo rápido e contínuo. Do uso exclusivo dos *mainframes* nos anos 70 e início dos 80, com os CPDs e seus técnicos centralizando informações e operações ao uso de redes, hoje tão comuns, com processamento distribuído entre PCs controlados diretamente por seus usuários, são tempos de mudanças contínuas.

Em Furlan (1994), encontra-se uma interessante analogia, em que as necessidades de informatização de uma empresa são comparadas à hierarquia das necessidades de Maslow. As necessidades de sistemas seriam atendidas, na seguinte ordem: 1º) nível operacional; 2º) suporte às operações do negócio; 3º) suporte à decisão; 4º) informações executivas; e 5º) informações industriais.

Sem entrar no detalhe de cada um dos níveis de sistemas, conforme anteriormente enumerado, pode-se afirmar que, na maioria das empresas, estão atendidos, pelo menos em parte, os dois primeiros níveis de necessidades de informatização, isto é, o operacional e suporte as operações do negócio.

Na atual fase da economia e política mundial, vive-se em uma sociedade em que o clima é de grandes mudanças. Para citar algumas, enumera-se: a) globalização; b) acirramento da competição; c) dinâmica acelerada da sociedade no tocante a padrões e opções de consumo e exigências crescentes por qualidade, inclusive ambiental. Em particular nas empresas, mudanças, tais como automação, crescimento do nível de informatização e formação de uma cultura de informática, são notáveis. Esta realidade vem exigindo esforços de atualização e adaptação contínuos. Do profissional de limpeza, que aprende a executar tarefas com um

Vaporetto, ao aposentado, que retira seu dinheiro com um cartão magnético em um caixa eletrônico, poucos mantiveram intactos seus hábitos e práticas nos últimos anos.

Nas empresas, estas mudanças refletem-se não somente em alterações no processo de desenvolvimento de produtos ou no processo produtivo, mas também nos demais processos. Não somente as áreas de Produção e *Marketing* precisam adaptar-se aos novos tempos, mas toda a empresa deve passar a trabalhar e ser gerida de outra forma. Existe uma demanda por mais agilidade, novas ferramentas, maior qualidade e precisão nas ações de todas as áreas, bem como por uma maior sensibilidade às influências do meio externo à empresa.

Neste contexto, vêm ganhando importância e atenção crescentes os processos de Planejamento Estratégico e a necessidade de informações rápidas, atualizadas, precisas e confiáveis, tanto sobre a empresa quanto sobre seu ambiente, para o subsídio aos procedimentos de gestão. Desta necessidade, do processo de informatização das empresas, da acessibilidade e de interfaces cada vez mais amigáveis dos computadores e programas, surgiram os Sistemas de Informações que vêm se tornando cada vez mais úteis e importantes.

O objetivo destes sistemas é satisfazer as necessidades de informações, eliminando intermediários entre a informação e seus usuários, disponibilizando-a de forma integrada e permitindo sua análise em vários níveis. Assim, espera-se que a partir de uma visão macro dos dados desejados o usuário possa aprofundar-se apenas naqueles que exijam maior atenção, seja para a definição de ações corretivas, ou para a identificação de oportunidades para a empresa.

Aqui não serão abordadas soluções voltadas a atividades operacionais da empresa, mas sim aquelas direcionadas para os níveis tático e estratégico das áreas de atuação do administrador. Neste sentido, identificam-se sistemas como SIG, DSS, EIS; e abordagens orientadas a bancos de dados, como *data warehouse*, *data mart*; e ainda as ferramentas de busca, como o *data mining*.

Esta multiplicidade de abordagens e ferramentas, aliada ao bombardeio promovido pela mídia, vendedores e consultores - que apresentam casos em que a tecnologia adotada em cada situação torna-se o fator essencial de sucesso - gera a necessidade de um levantamento mais abrangente do que vem sendo proposto, com o objetivo de organizar e sistematizar informações sobre o assunto, facilitando a análise e comparação das várias alternativas. Percebe-se, também, uma noção de urgência na tomada de atitudes, uma idéia de que aqueles que não investirem agora sofrerão as conseqüências da falta de subsídios, de controle sobre a situação. Esta noção acaba sendo transmitida independente da análise de cada caso ou proposta específicos, o que nem sempre corresponde à realidade.

Diante desse contexto, através do presente trabalho, faz-se uma definição e caracterização das soluções de *data warehousing* e elabora-se um modelo de *data mart* para a área de informática do Tribunal Regional do Trabalho - 12ª Região. No desenvolvimento deste modelo, são utilizadas soluções e tecnologias que posteriormente serão empregadas na criação de *data marts* direcionados a outras áreas do Tribunal, sendo a área de informática escolhida para este primeiro projeto pela facilidade de acesso às informações e aos usuários e pelo tamanho do modelo - coerente com os propósitos deste trabalho.

A motivação para o desenvolvimento deste trabalho parte do reconhecimento da necessidade de esclarecimento e contextualização dos conceitos e ferramentas de *data warehousing* em ascensão na área de informática, sendo, portanto visto como uma oportunidade de sistematizar e empregar soluções importantes apresentadas em publicações de informática e de administração.

1.1 Apresentação do Tema-Problema

Atualmente, a necessidade e a importância da informatização são amplamente reconhecidas no Tribunal, sendo esta informatização um processo

contínuo de aplicação das tecnologias e recursos que vão sendo desenvolvidos e disponibilizados para o atendimento às diversas necessidades e atividades.

Com o atendimento às necessidades operacionais básicas, vem crescendo a demanda por sistemas voltados à administração do Tribunal. Os administradores já viram os resultados da informatização das atividades operacionais e sentem a necessidade de acesso diferenciado às informações. Já não são suficientes as consultas e relatórios existentes. Mais e mais é preciso aplicações que ofereçam recursos mais ágeis e precisos para análises e planejamento. Neste contexto, surge a necessidade do estudo e definição de uma solução voltada a estas necessidades. Com este propósito, foi observado que, com a evolução do processo de informatização, o processo de planejamento de informática vem tornando-se cada vez mais complexo, apresentando-se como um modelo interessante para o desenvolvimento deste trabalho.

Inicialmente, quando nenhuma área estava atendida, os resultados de qualquer ação eram mais visíveis, sendo mais fácil detectar prioridades de investimento. Mas, na medida em que a informatização vai se concretizando a tendência é aparecerem distorções, como: máquinas superdimensionadas em atividades não tão importantes por questões de visibilidade ou acesso do usuário à administração; e tarefas importantes na atividade fim da empresa sendo executadas em máquinas inadequadas, exigindo muito mais tempo e esforço.

Para uma melhor aplicação dos recursos disponíveis, torna-se cada vez mais importante a disponibilidade de informações sobre o processo de informatização, tais como, a distribuição dos equipamentos e sistemas, as solicitações pendentes e atendidas e os projetos em andamento.

No TRT, o controle dos equipamentos de informática é feito pelo Serviço de Material e Patrimônio (SEMAP). Os computadores, impressoras, e demais equipamentos são tratados como material permanente, da mesma forma que mesas, cadeiras ou condicionadores de ar. Quanto aos *softwares*, são adquiridos como serviços, sendo sua instalação e uso controlados apenas pelo SEINFO. Além

dos equipamentos e sistemas adquiridos, o SEINFO precisa de informações sobre as necessidades e o uso de sistemas desenvolvidos internamente.

Assim, para o bom desempenho de suas atividades, os administradores da área de informática do Tribunal necessitam, para a tomada de decisões, de informações sobre:

- ◊ os equipamentos (*hardware*), sua configuração e localização;
- ◊ os *softwares* (word, excel, fácil jurisprudência,...) instalados;
- ◊ os sistemas desenvolvidos pelo Tribunal;
- ◊ os *softwares* básicos (rede, sistema operacional, banco de dados,...); e
- ◊ as solicitações de instalação e atualização de *software* e *hardware*.

O que se observa, no entanto, é que os sistemas utilizados para controle do patrimônio do Tribunal não atendem às necessidades específicas de informação da área de informática, por não reunirem dados sobre *hardware* e *software*, nem disporem de dados históricos sobre o processo de informatização.

Nesse contexto, define-se como tema-problema para o atual trabalho, o estudo e elaboração de um *data mart* para atender a necessidades de informação da área de informática do Tribunal Regional do Trabalho - 12 Região.

1.2 Justificativa

1.2.1 Importância

Uma solução que agrupe as informações relativas à informática no TRT é importante principalmente para os gestores da área que precisam de uma visão global da situação da área e do que vem sendo feito, onde têm sido aplicados os recursos, além da identificação das áreas que estão sendo priorizadas na distribuição de equipamentos e sistemas. É importante que estas informações estejam disponíveis de forma agrupada, para análise global da situação de uso dos recursos de *hardware* e de *software* em cada área. Com estes dados, torna-se

possível uma melhor identificação das situações mais críticas. Também, facilita a negociação com as demais áreas e com a Presidência e a resposta a solicitações e elaboração de justificativas de pedidos de compra e orçamento.

Além da gerência da área, os setores de administração de sistemas operacionais e bancos de dados também serão beneficiados com a disponibilidade de uma base de dados única e atualizada, contendo todas as informações sobre as características dos equipamentos e sistemas. Estas informações são necessárias para o planejamento e execução das atividades destes setores e, atualmente, são coletadas e mantidas por cada setor conforme a demanda.

1.2.2 Oportunidade

Atualmente, o Tribunal está em processo de padronização das ferramentas utilizadas, tanto nos equipamentos servidores, quanto nas estações. Incluindo, também, as ferramentas de desenvolvimento de sistemas, e os *softwares* adquiridos. Assim, pretende-se a adoção de ferramentas padrão, como por exemplo: Word para edição de textos, Excel para aplicações de cálculos, Oracle Developer 2000 e Visual Basic para o desenvolvimento de sistemas, e Oracle Server como banco de dados.

Além da oportunidade de agilizar e facilitar o trabalho na área de informática, a implantação deste projeto permitirá o domínio de uma tecnologia pela qual já existe demanda em outras áreas como Material e Patrimônio e Controle de Processos.

Também contribui para este processo de padronização de ferramentas, a ênfase da administração na integração das diversas áreas do Tribunal e JCs e na disponibilização de recursos através da *internet* e *intranet*, bem como a redução da impressão de relatórios e tramitação de papéis. Neste contexto, obtém-se o domínio da tecnologia de desenvolvimento e manutenção de *data warehouses*,

principalmente se os dados forem disponibilizados para consultas através da intranet, o que se torna especialmente interessante.

1.2.3 Viabilidade

O desenvolvimento deste projeto, ou seja, a definição de um modelo de *data mart* para armazenar informações sobre a informatização do Tribunal tem sua execução favorecida pelo estado da informatização do mesmo. Grande parte dos dados necessários já foram analisados e estão sendo armazenados em sistemas em produção. O projeto pode ser implementado utilizando-se as ferramentas de desenvolvimento de sistemas existentes, não implicando em novas aquisições ou entrando em conflito com as políticas adotadas. Também contribui para a viabilidade deste, o interesse e a necessidade do domínio das tecnologias de *data warehouse* pela área de informática. Como citado no item sobre a oportunidade do projeto, já existe demanda neste sentido. A escolha da área de informática como alvo do projeto também é conveniente por proporcionar a definição de um modelo pequeno, com poucos usuários, sendo estes de fácil acesso e diálogo por sua familiaridade com as tecnologias e resultados esperados, viabilizando a definição de um projeto de dimensões compatíveis com os recursos e prazos disponíveis.

2. OBJETIVOS

A seguir são apresentados os objetivos geral e específicos deste trabalho.

2.1 Objetivo geral

Esclarecer e contextualizar conceitos e ferramentas de *data warehousing*, sistematizando e empregando estes conceitos e métodos. Definição de um modelo adequado de um *data mart* para suprimento de necessidades de informação da área de informática do Tribunal Regional do Trabalho - 12 Região, reunindo dados sobre equipamentos e *softwares* utilizados pelo Tribunal.

2.2 Objetivos específicos

- Estudar e analisar tecnologias disponíveis para suprimento de informações gerenciais, com ênfase em *data warehousing*, área em que se enquadra o *data mart*.
- Determinar os processos executados pela área de informática do TRT.
- Selecionar os processos da referida área que serão modelados pelo sistema.
- Definir o nível de detalhe em que os processos selecionados serão descritos.
- Especificar as dimensões dos processos modelados, ou seja, as relações entre os dados que serão armazenados.

- Identificar os fatos a registrar, expressando-os no nível de detalhe em que os processos foram definidos.
- Analisar as dimensões definidas para os processos, estabelecendo descrições apropriadas e significativas para os usuários.
- Verificar as modificações das dimensões referidas, prevendo-as para garantir a manutenção e consistência dos dados históricos.
- Definir a necessidade ou a conveniência de armazenar dados agregados e outras decisões decorrentes do tamanho do *data mart* e das consultas esperadas.
- Determinar a amplitude do histórico a ser mantido no *data mart*.
- Definir a periodicidade de atualização dos dados armazenados no *data mart*.
- Definir a ferramenta de consulta para recuperação dos dados no *data mart*.

3. FUNDAMENTAÇÃO TEÓRICA

Trata-se da revisão da literatura existente quanto aos assuntos abordados no estudo, especialmente no que tange a teoria de tecnologia de informática e *data warehousing*.

3.1 A crescente necessidade de planejamento e informação

Qualidade total, reengenharia, administração por objetivos, virtualização, organização em processo de aprendizagem, corporações achatadas, organizações caólicas... é grande o número de soluções para tornarem as empresas mais competitivas nos tempos atuais, de grandes e contínuas mudanças. Kotler (1991) afirma que nos anos 50 e 60 (nos EUA) era suficiente para os administradores o uso do planejamento operacional e, ainda assim, era pouco provável que mesmo um administrador inexperiente não fosse bem sucedido. Hoje em dia as coisas estão bem diferentes e, pior, não estão estáveis. Houve as crises do petróleo; a globalização da economia; a invasão dos mercados pelos produtos japoneses - mais baratos e com qualidade; esquemas de proteção estão sendo abandonados com a criação dos Mercados Comuns como o MERCOSUL, o NAFTA e o Mercado Comum Europeu; estão sendo criadas restrições a produtos que não atendam a padrões de qualidade como as normas ISO e, aparentemente, tem-se mais novidades ainda por vir.

Nesta nova realidade, surge a necessidade de adaptação e resposta às mudanças constantes, ocasionando a demanda por um novo processo de planejamento administrativo que proporcione às empresas condições de se manter em uma boa posição, apesar de problemas ocorridos em qualquer um de seus

negócios ou linhas de produtos. Neste sentido, surgiu e vem sendo utilizado o Planejamento Estratégico.

O Planejamento Estratégico, segundo Kotler (1991), tem três idéias básicas, como segue:

- a) uso pela empresa de um portfólio de investimentos avaliados continuamente quanto à conveniência de início, manutenção, desaceleração ou encerramento de cada investimento;
- b) avaliação do potencial de lucro futuro de cada investimento; e
- c) utilização de estratégias, ou seja, para cada tipo de negócio, a empresa deve definir uma estratégia para alcançar seus objetivos de longo prazo, com base na análise de seus pontos fortes e fracos e das oportunidades e ameaças oriundas do mercado.

Em Stoner (1985), encontra-se que:

As escolas clássica, comportamentista e quantitativa de pensamento administrativo concentram-se em aspectos da organização que os administradores podem influenciar diretamente. Estas escolas de administração disseram aos administradores quantos subordinados eles deveriam ter, por que é importante melhorar o ambiente de trabalho e como eles poderiam usar o computador para facilitar sua tomada de decisões. Mas, em sua preocupação com o ambiente interno das organizações, subestimaram a importância do ambiente externo. O clima político da sociedade e o que as pessoas de fora da organização achavam dela nunca foram o principal alvo de interesse das três escolas de administração (...) Se o ambiente externo das organizações for razoavelmente estável e previsível, haverá menos necessidade de os administradores se preocuparem diretamente com ele. Em nossa época, porém, o ambiente externo está passando por rápidas mudanças que tem efeitos importantes e, às vezes, bastante imprevisíveis sobre as organizações e sua administração. (1985, p. 41)

Ainda sobre a influência do ambiente sobre as organizações, observa-se que um sistema é fechado quando não sofre influência das variáveis

externas, sendo determinístico, e não probabilístico. Se as organizações pudessem ser consideradas como sistemas fechados, a eficácia organizacional seria obtida e mantida através do controle das variáveis organizacionais dentro de determinados limites, ou seja, poder-se-ia definir um conjunto de regras para a manutenção de relações adequadas entre as partes da organização que proporcionariam o sucesso. Entretanto, as organizações são sistemas abertos, sofrendo o impacto contínuo das mudanças em seu ambiente (variáveis externas). Desta forma, o ambiente da organização não possui limites definidos e contém um grande número de variáveis, que não podem ser controladas pela mesma.

As organizações podem ser vistas como sistemas dentro de sistemas, dado que a sociedade, com seus vários grupos, organizações e movimentos pode ser considerada um grande sistema. Cada organização recebe parte dos valores dominantes do seu ambiente. Mas, ao mesmo tempo, seus membros transitam pelo ambiente externo à empresa, e possuem influência sobre a estrutura social e cultural da sociedade. Os membros de uma organização de trabalho são, ao mesmo tempo, membros de muitos outros grupos, competidores entre si ou não. (Chiavenato, 1987)

Como se percebe, não há fórmula ou estratégia capaz de propiciar o sucesso à todas as empresas concorrentes. Mas uma coisa pode ser considerada certa: a necessidade de reconhecimento e adaptação às constantes mudanças do mercado e do ambiente.

Como exemplo, pode-se citar Contador (1995), que a partir de uma análise dos acontecimentos recentes e da atualidade, afirma que o comando das organizações tem sido exercido pelas seguintes áreas:

- ◊ anos 50 - área financeira;
- ◊ anos 60 - planejamento de longo prazo (finanças e *marketing*);
- ◊ anos 70 - *marketing*;
- ◊ anos 80 - coordenação - *marketing*;
- ◊ anos 90 - manufatura.

Por que o comando da manufatura?

O autor afirma que a organização deve envolver todas as áreas sob o comando da manufatura, por ser esta a responsável por:

- ◊ qualidade;
- ◊ baixo custo;
- ◊ menor prazo de entrega;
- ◊ flexibilidade para a troca de produtos;
- ◊ diversificação de produtos;
- ◊ lançamento de novos produtos;

Enfim, é justamente a área da manufatura que vai proporcionar a agilidade e a competitividade da empresa.

Neste sentido, a empresa deve traçar uma estratégia global para a consecução de seus objetivos, a partir da análise da cadeia produtiva de forma integrada, e não por setor ou área, a fim de que todos os setores trabalhem em uma mesma direção, maximizando os resultados obtidos.

Sobre as mudanças recentes da administração e da sociedade, serão reproduzidos a seguir, três trechos do artigo “A Administração, essa desconhecida”, de Peter F. Drucker (1988).

Durante os anos 20 e 30, o “planejamento cresceu além dos gráficos de Gantt criados em 1917 e 1918 para planejar a produção de materiais de guerra, e assim também o uso de lógica analítica e estatística, a qual se valeu da quantificação para converter experiência e intuição em informação, definições e diagnósticos” (Drucker, 1988, p. 64). Drucker (1988) ainda continua:

Poucas instituições na História da humanidade emergiram tão rápido quanto a administração ou tiveram um impacto tão grande quanto surpreendente. Em menos de 150 anos, a administração transformou a estrutura social e econômica dos países desenvolvidos. Ela criou uma economia global e estabeleceu novas regras para os países que participassem dessa economia de igual para igual. E a própria administração também se transformou. Para ser mais exato, o dever fundamental da administração continua o mesmo: tornar possível às pessoas o

trabalho conjunto, dando a elas objetivos comuns, valores comuns, a estrutura certa e o treinamento e desenvolvimento que elas precisam para reagir e atuar sobre as mudanças. Mas o sentido principal desta tarefa mudou. Até porque o desempenho da administração transformou a mão-de-obra composta em sua maioria de trabalhadores não-especializados numa mão-de-obra com alto conhecimento e educação. (...)

Até a Primeira Guerra Mundial, era axiomático que levaria um longo tempo (Adam Smith disse centenas de anos) para um país ou região desenvolver uma tradição de trabalho e de especialização em habilitações manuais e organizacionais necessárias para produzir e por no mercado um produto qualquer, seja ele um pano de algodão, seja um violino. Mas, durante a Primeira Guerra Mundial, grande número de pessoas completamente sem especialização, pré-industriais, teve que se transformar em trabalhadores produtivos em praticamente algumas horas. (Drucker, 1988, p. 60-61)

Neste contexto de mudanças, os administradores e as organizações respondem ao ambiente externo procurando prever as tendências do ambiente, e influenciar ou se adaptar, conforme o que for possível, ou mais conveniente.

Na medida em que a competição acentua-se, em praticamente todos os mercados os executivos passam a se interessar mais por dados externos do que pelos dados operacionais da empresa. Cada vez mais, tornam-se importantes as informações sobre o ramo de atividades da empresa, contemplando dados sobre o meio organizacional em que esta está inserida, incluindo: concorrentes, governo, consumidores e tendências de mercado. A comparação do desempenho relativo aos concorrentes ou o crescimento comparado ao do mercado e da economia como um todo, passam a ser mais significativos do que os números internos isolados do contexto. (Furlan, 1994)

Os executivos devem estar cientes de que as necessidades e a importância das informações podem crescer (em volume e frequência) de maneira exponencial em relação ao crescimento das empresas. Entretanto, para que esta evolução ocorra, as informações precisam ser confiáveis, relevantes e disponíveis em tempo, para viabilizar decisões corretas aos executivos. Para tal, deve ser

considerado o valor da informação, que depende do impacto que esta provoca nas decisões dos executivos, e da sua utilidade, considerando o tempo de utilização desta pela empresa. (Oliveira, 1993)

Com o objetivo de disponibilizar informações de forma adequada, e de fornecer aos gestores uma ferramenta voltada à identificação de dados estratégicos e aspectos críticos das atividades, oferecendo, desta forma, informações relevantes para o processo de administração das empresas, vêm sendo desenvolvidas ferramentas e tecnologias voltadas ao armazenamento e recuperação de informações gerenciais, objeto deste trabalho. Estes programas procuram ajudar o gestor a identificar o momento certo de tomar a decisão e provê-lo de subsídios para a decisão mais acertada. “Na empresa, ao contrário do que ocorre nos filmes de faroeste, nem sempre sacar a arma primeiro significa vencer o duelo com o concorrente. O que importa, mesmo, é acertar o alvo: garantir o crescimento e a continuidade da organização.” (Machado, 1996, p. 49)

3.2 A tecnologia de informação como apoio à decisão e distribuição de informações

Antes de seguir tratando sobre os sistemas em questão, são analisados alguns conceitos básicos, referentes à área de sistemas:

“Sistema é um conjunto de partes interagentes e interdependentes que, conjuntamente, formam um todo unitário com determinado objetivo e efetuam determinada função” (Oliveira, 1993, p. 23). Ainda, “ambiente de um sistema é o conjunto de elementos que não pertencem ao sistema, mas qualquer alteração no sistema pode mudar ou alterar os seus elementos e qualquer alteração nos seus elementos pode mudar ou alterar o sistema” (Oliveira, 1993, p. 25). Nesse contexto, o ambiente externo compreende os elementos de fora da organização que são importantes para seu funcionamento. As organizações não são auto-suficientes, nem isoladas. Efetuam trocas de recursos com o ambiente externo e dele dependem: matérias primas, dinheiro, mão de obra e energia são insumos do ambiente externo

adquiridos pelas organizações, que os transformam em produtos ou serviços e os oferecem ao ambiente externo. (Stoner, 1985)

Já, “informação é o dado trabalhado que permite ao executivo tomar decisões” (Oliveira, 1993, p. 34). Vale lembrar que, “gerencial é o processo administrativo (planejamento, organização, direção e controle) voltado para resultados” (Oliveira, 1993, p. 37). Assim, para elevar a qualidade do seu desempenho tem-se usado várias ferramentas. Entre elas, destacam-se os SIG. “Sistema de Informações Gerenciais (SIG) é o processo de transformação de dados em informações que são utilizadas na estrutura decisória da empresa, bem como proporcionam a sustentação administrativa para otimizar os resultados esperados” (Oliveira, 1993, p. 39).

Alguns autores classificam os sistemas voltados às gerências em: Sistemas de Informação Executiva - EIS - dirigido à alta administração; e Sistemas de Suporte à Decisão/Sistemas de Informações Gerenciais - DSS/SIG - para a média administração. Conforme observado por Machado (1996), atualmente, as transformações nos perfis dos sistemas utilizados nas empresas (movidas pelas mudanças tanto tecnológicas quanto administrativas a partir dos processos de redução de níveis hierárquicos, enriquecimento de tarefas, entre outros), fizeram com que grande parte das distinções entre SIG, EIS e DSS fossem eliminadas, restando os sistemas de nível operacional - tais como contabilidade e folha de pagamento - conhecidos como OLTP (*On Line Transacion Processing*); e os sistemas voltados à administração - por sua vez conhecidos como OLAP (*OnLine Analytical Processing*).

As maiores fontes de dados para a análise e tomada de decisão de que a empresa dispõe são os sistemas desenvolvidos para o suporte às atividades operacionais da empresa, tais como os sistemas de controle de pessoal, vendas, produção e estoques. Entretanto, as necessidades e características dos sistemas operacionais e de análise são bastante diferentes. Nos sistemas operacionais, a demanda maior é por performance e informação atualizada, não havendo a

necessidade de manutenção dos dados históricos, que se presentes, acarretam em sobrecarga e lentidão do sistema nos processos de inclusão, consulta e recuperação de dados. Outra característica dos sistemas operacionais é que as operações executadas são padronizadas, seguindo procedimentos-padrão da empresa. Por sua vez, os sistemas voltados à análise dependem basicamente de dados históricos, para efetuar operações de totalização, agrupamento, comparação e o que mais for sendo requerido a partir da análise de cada um dos resultados obtidos e dos interesses do analista. Em poucas palavras, enquanto os sistemas operacionais demandam performance e possuem operações padronizadas, os sistemas para análise demandam flexibilidade e abrangência, com uma maior tolerância a tempos de resposta maiores.

Destas diferenças, surgiu a teoria que sustenta os *data warehouses*, que parte da idéia de separar em bancos de dados distintos os dados e usuários de sistemas operacionais e de análise.

Definindo-se bancos de dados, tem-se que:

Algumas fontes asseguram aos leitores que qualquer conjunto de informações, as quais tenham um motivo para serem agrupadas, como, por exemplo, um arquivo comum de correspondência ou uma lista de números de telefone, é um banco de dados. O meu dicionário on line (American Heritage) transforma o termo "database" em duas palavras (data base) e prossegue definindo-o num contexto moderno, como "uma coleção de dados, os quais são dispostos de forma que possam ser buscados pelo computador. (Holtz, 1994)

A partir da definição de bancos de dados, tem-se que *data warehousing* é uma tecnologia em que é empregado um banco de dados voltado para o suporte à decisão de usuários finais, tendo seus dados derivados dos bancos de dados operacionais. Neste, são armazenados grande volume de dados e históricos, para permitir a análise de operações passadas e prever situações futuras. (Dantas e Filho, 1997)

A influência do ambiente externo sobre as empresas e os administradores varia de acordo com o tipo e a finalidade da organização, varia também entre os cargos e funções e entre os níveis hierárquicos desta. Esta variação de influência também existe entre os funcionários responsáveis por providências claras e os gerentes de venda de divisões, e sobre os executivos mais do que sobre os empregados de escritório. (Stoner, 1985)

A grande vantagem do uso de sistemas integrados está na visão global que estes proporcionam. Eles exigem cuidado em seu desenvolvimento e implantação, mas, se adequados e bem ajustados, podem significar o diferencial tecnológico em relação à concorrência. Implantar um desses programas significa mudar a cultura da empresa, o que não é coisa que se faça de uma hora para a outra.

Sobre a importância do SIG, em Oliveira (1993) encontra-se: “se você conhece o inimigo e conhece a si mesmo, não precisa temer o resultado de cem batalhas. Se você se conhece, mas não conhece o inimigo, para cada vitória ganha, sofrerá também uma derrota. Se você não conhece nem o inimigo nem a si mesmo, perderá todas as batalhas”. (Sun Tzu apud Oliveira, 1993, p. 15)

Deve-se ter em mente que as inovações administrativas devem representar a “aplicação de conhecimento no trabalho, a substituição de conjecturas, músculos e labuta por sistemas e informação.” (Drucker , 1988, p. 64)

Desde 1945, quando a primeira bomba atômica explodiu e o primeiro computador começou a funcionar, “ambos os modelos de tecnologia e organização se tornaram um modelo biológico-interdependente, intensamente especializado e organizado pelo fluxo de informação.” (Drucker , 1988, p. 65)

Hoje em dia, a resposta de uma organização ao seu ambiente pode ser crítica para sua existência e, neste contexto, é necessário que os administradores disponham de um método ou orientação para manter e melhorar seu desempenho neste ambiente em transformação. (Stoner, 1985)

O uso de computadores, em vez de papéis, na administração das empresas não é somente uma questão de modernidade, mas de agilidade e precisão. Com a complexidade das operações e a competição no mercado, as empresas estão se voltando para dinamizar seu processo de decisão. Com um SIG é possível o acompanhamento diário de resultados, tabulando dados das diversas áreas da empresa e depois exibindo-os de forma gráfica e simplificada. (Furlan, 1994)

Para que cumpra seus objetivos, o SIG deve estar dirigido aos negócios da empresa e, portanto, “consolidar uma inteligência para negócios, que corresponde a um processo analítico de transformação de dados desagregados de mercado, concorrência, produtos, consumidores e tecnologia, entre outros fatores importantes em estratégias operacionalizáveis a respeito dos concorrentes e do desempenho, capacitação e interações da empresa.” (Tyson apud Oliveira, 1993, p. 86)

Os SIG devem atuar como elementos polarizadores dos eventos empresariais provenientes das atividades, tanto internas como externas à empresa (Oliveira, 1993). Assim, o executivo tem condições de identificar, de forma imediata, os fatores críticos de sucesso, conforme critérios determinados por ele mesmo, permitindo que este concentre suas atenções nas situações críticas; evitando a perda de tempo com situações de menor importância; e assim, tornando possível a concentração de esforços na análise dos indicadores, para a adoção de ações corretivas. (Furlan, 1994)

Um SIG inteligente, segundo Oliveira (1993, p. 86) deve ser capaz de:

- ◊ “evitar surpresas para a empresa;
- ◊ identificar oportunidades e ameaças ambientais;
- ◊ consolidar vantagem competitiva real; e
- ◊ facilitar os planejamentos de curto, médio e longo prazos.”

Os principais benefícios da implementação de um Sistema de Informação Executiva - EIS, são, segundo Furlan (1994, p. 15):

- ◊ “tomada de decisões oportuna e efetiva;
- ◊ redução no tempo de obtenção de informações;
- ◊ melhoria e agilização da comunicação interna entre áreas e executivos;
- ◊ esforços dirigidos para a consecução de objetivos.”

O que se observa é que, a não agilidade nas decisões, ocasionada pela dificuldade de manipulação das informações corporativas, pode implicar na perda da vantagem competitiva, gerando dissabores e custos adicionais, pelo fato do nível executivo não possuir uma ferramenta adequada à execução de suas atividades. (Furlan, 1994)

Entretanto, a adoção de sistemas para o suporte a decisão não pode ser considerada uma panacéia, a solução para todos os problemas, uma vez que,

Uma parte dos investimentos equivocados em tecnologia tem origem na corrida alucinada das corporações para arrancar na frente da concorrência em todos os campos - inclusive o tecnológico. Na duríssima era da globalização, as empresas são impelidas pelo espírito de sobrevivência a ser mais rápidas, mais leves, mais flexíveis, mais produtivas, mais lucrativas. Nesse clima darwinista, como resistir ao apelo do computador? Como distinguir o modismo tecnológico oco dos grandes aliados tecnológicos que jogam o mundo dos negócios para um patamar superior de eficiência? (Militello, 1997, p. 40)

Nem a mais recente novidade, nem a solução do concorrente são a melhor solução para a empresa, pois “máquinas idênticas com *softwares* idênticos executarão coisas maravilhosas em companhias bem-estruturadas e organizadas, mas serão apenas uma fonte de gastos em outras com administração inferior.” (Militello, 1997, p. 40)

Muitas vezes, nos meios empresariais, age-se como se os computadores possuíssem poderes mágicos para resolver problemas de gestão, racionalizar processos e aumentar a produtividade. Não se discute como estes serão recebidos dentro da empresa, nem se procura provar como eles aumentarão os lucros. Em vez disso, impera o medo da concorrência (Gurovitz, 1997). Parece ser opinião generalizada que, “se seu inimigo tem metralhadoras, você também tem que ter metralhadoras. Mas não são metralhadoras que ganham a guerra. São pessoas. Você só vence servindo os consumidores e gerando valor. Não imitando o que fazem os outros.” (Gurovitz, 1997, p. 88)

Como exemplo, é interessante observar o seguinte caso: num depósito do Pão de Açúcar,

Cada homem movimentava 40 caixas em uma hora. Nos EUA, esse número chegava a 150. Depois de viajar ao Canadá e trazer de lá um sistema mais racional de trabalho, o que incluiu a informatização, o chefe do depósito diz já movimentar 100 caixas e quer chegar a 180. Mas será que só a tecnologia foi responsável pela mudança? Não. O método de trabalho no depósito estava errado. Para a coisa começar a funcionar melhor, o essencial foi adotar duas docas e rearranjar as caixas. Isso poderia ter sido feito sem nenhum computador. (Gurovitz, 1997, p. 92)

Em Degen (1989) tem-se que “competitividade é a base do sucesso ou fracasso de um negócio em mercados onde há livre concorrência ... competitividade é a correta adequação das atividades do negócio ao seu microambiente. Esta adequação correta consiste no bom atendimento das necessidades dos clientes, diferenciação nesse atendimento em relação aos concorrentes, bom relacionamento com os fornecedores e a existência de barreiras à entrada de novos concorrentes.” (Degen, 1989 p.106)

O empreendedor deve preocupar-se em definir a estratégia para maximizar a adequação do negócio ao seu macroambiente e, desta forma, obter vantagens em relação aos seus concorrentes (atuais e potenciais).

Como pode ser facilmente percebido, ações isoladas em setores da empresa proporcionam efeitos bem inferiores àqueles que seriam obtidos se toda a organização se engajasse na execução dos objetivos propostos. Neste sentido, a empresa deve munir-se das ferramentas que melhor se adequarem à sua realidade e necessidades. Este processo deve ser contínuo sendo que as estratégias e mesmo os objetivos propostos devem ser permanentemente avaliados quanto ao alcance dos resultados estimados e a viabilidade de seqüência do plano. O plano deve refletir as metas e estratégias atuais e futuras da empresa, não podendo ficar restrito ao período de sua elaboração, se houverem alterações importantes na conjuntura ou nas condições da empresa este deve passar a refleti-las a fim de que possa ser útil à administração e que as metas possam ser alcançadas.

Como se sabe, no cenário atual em que as empresas reagem mais rapidamente às mudanças no mercado e, ao mesmo tempo, reduzem suas camadas intermediárias de gerência, as informações tornam-se cada vez mais críticas ao próprio negócio (Filho & Dantas, 1997). É necessário construir mecanismos que circulem informações vitais por toda a organização, permitindo a tomada de decisões mais rápidas e certas. Neste contexto, alguns segmentos de indústrias, como o varejo, bancos, seguros e empresas aéreas e de telecomunicações, fornecem exemplos típicos de obtenção de vantagens competitivas com a aplicação de um *data warehouse*. Assim, em breve, essa tecnologia vai se tornar uma ferramenta estratégica para a maioria das organizações.

Em relação a um planejamento estratégico da informação, historicamente, a maioria das empresas no Brasil, principalmente na informática, relegam este planejamento estratégico a último plano. Uma minoria voltou-se ao planejamento de necessidades tecnológicas apenas (os assim chamados PDIs), como se o negócio principal da empresa fosse a informática. Um grupo muito restrito voltou-se realmente ao planejamento estratégico da informação, em que os planos da informática são atrelados aos planos empresariais. Um erro grave, pois os objetivos deveriam ser de dar suporte à informação gerencial.

Concluindo, o que se observa é que mudanças relevantes relacionadas à forma como os negócios são direcionados e conduzidos, à maneira como as pessoas se organizam em busca de objetivos comuns e ao uso de novas tecnologias como meio de otimização dos processos do negócio não podem ser consideradas modismos, mas sim estratégias de sobrevivência. Com grande responsabilidade nessas mudanças, a tecnologia da informação é o caminho encontrado pelas organizações para viabilizar a execução de novas estratégias de negócios, estratégias com ênfase no mercado, maior aproximação ao cliente, qualidade em serviços e crescimento planejado.

Assim, o potencial apresentado pela tecnologia da informação torna possível a estruturação de uma empresa orientada a processos estratégicos, disponibilizando a informação onde a decisão deve realmente ser tomada. Essa nova estrutura, necessária e indispensável, deve substituir a concorrência interna pela colaboração e garantir o comprometimento global com os objetivos corporativos.

Ainda hoje, a disponibilidade da informação no momento crítico de decisão constitui-se em um forte diferencial competitivo, mas se sabe que, em um futuro muito próximo, a disponibilidade será necessidade básica, e o grande diferencial estará no modelo decisório.

E o preço? Será sempre tão caro (no mínimo três milhões de dólares para a implementação)? Definindo com clareza os objetivos de se implementar um projeto de *data warehouse*, utilizando-se recursos existentes (PCs e um *software* OLAP) e escolhendo uma área realmente carente de informações para ser o alvo do projeto, pode ser diminuído em muito o preço dos projetos de *data warehouse*, tornando-se viável também em empresas de médio porte. (Dantas e Filho, 1997)

3.3 Fontes e formas de recuperação de informações gerenciais

Na história do desenvolvimento de sistemas, primeiramente, os esforços foram feitos no âmbito dos sistemas operacionais e dados que estes

processam. Entretanto, como já citado, não é prático nem conveniente manter dados nos sistemas operacionais indefinidamente. O desenvolvimento de uma estrutura com a finalidade de arquivar os dados processados pelos sistemas operacionais para análises posteriores surgiu somente depois, a partir das facilidades providas pelas tecnologias de banco de dados e cliente-servidor. A convivência entre sistemas operacionais e de análise, utilizando os mesmos bancos de dados, é inadequada porque os requisitos fundamentais destes sistemas são diferentes: os sistemas operacionais demandam performance, enquanto os de análise demandam flexibilidade e abrangência. Raramente é aceitável que aplicações de análise do negócio degradem a performance do sistema operacional. (Gupta, 1997)

Neste sentido foram sendo desenvolvidas soluções com a finalidade atender às necessidades de análise de informações, como monitorar e comparar as operações atuais com as passadas, prever situações futuras, permitindo a execução de análises estratégicas eficazes, e de análises de tendências, sem degradar os sistemas operacionais da empresa.

Os itens que seguem descrevem as principais fontes de dados para o suporte a decisão dentro das empresas, caracterizando a evolução das formas de armazenamento, recuperação e tratamento destas informações.

Legacy systems

Legacy systems compreende o legado de sistemas da empresa, desenvolvidos em sua fase inicial de informatização. Geralmente escritos em linguagem de programação Cobol, os *legacy systems* acabaram se tornando difíceis de abandonar, pois com o tempo estes sistemas foram se desenvolvendo e aperfeiçoando a ponto de abranger uma grande quantidade de conhecimento e regras do negócio que são incrivelmente difíceis de serem levados a uma nova plataforma de sistema ou aplicação. (Gupta, 1997)

Estes sistemas ainda são a maior fonte de dados para análise, entretanto, o tempo de recuperação de informações e desenvolvimento de relatórios é bastante grande, dependendo continuamente de pessoal técnico para geração e manutenção.

Informações geradas ou extraídas em estações de trabalho (*desktop*)

Nos anos 80, a disseminação dos microcomputadores (PCs) proporcionou várias opções e oportunidades para análises de negócios. O espaço entre o programador e o usuário foi diminuindo na medida em que os analistas de negócios passaram a ter acesso a ferramentas como planilhas para análise e representação gráfica, havendo usuários que criaram bancos de dados em seus computadores para armazenar e trabalhar sobre dados extraídos dos grandes sistemas operacionais (*legacy sources*). (Gupta, 1997)

O inconveniente desta prática, é que os dados tornam-se fragmentados e orientados a necessidades muito específicas. Cada usuário individual obtém somente a informação que deseja. Não sendo padronizados, estes extratos de informação não são capazes de atender a usuários e necessidades diversas. Desta forma, o tempo e o custo envolvidos no atendimento a necessidades de um único usuário acabou por se mostrar proibitivo. Além de demandar tempo e esforços do pessoal técnico de informática para gerar e converter os extratos dos dados das *legacy sources* necessários a cada usuário, este método de trabalho assume que o usuário final tem tempo para gastar no gerenciamento de dados em planilhas, arquivos e bancos de dados, sendo que, dada a oportunidade de escolha, a maioria deles julga mais eficiente dedicar-se à análise e ferramentas de análise, que ao gerenciamento de dados. (Gupta, 1997)

Sistemas de Informações Gerenciais

Por muito tempo, as organizações têm implementado seus sistemas de suporte à decisão em um ambiente projetado para suportar as atividades

operacionais da empresa. Neste ambiente, as decisões operacionais que são aquelas direcionadas para o dia a dia são satisfatoriamente suportadas por meio de consultas pré-definidas às bases de dados e relatórios. (Tecnologias, 1999) Conforme o porte e o grau de informatização da empresa, esta abordagem já é uma grande evolução. Em Grego (1998), encontra-se que:

Administrar uma microempresa numa época de turbulência econômica como esta não está mole. Quais são os produtos mais rentáveis? Há algum serviço que deve ser cortado porque não dá lucro? Quais custos podem ser reduzidos? Onde estão as oportunidades para novos negócios? Um bom programa administrativo deve ajudar o profissional a se desvencilhar da burocracia e ainda facilitar a tomada de decisões como estas. Em suas versões mais recentes os aplicativos para pequenas empresas estão mais flexíveis e trazem muitos dados para ajudar nas decisões.

Entretanto, a partir do uso destes sistemas, os gerentes passam a sentir necessidades de outras informações e formas de consultas. Com o tempo, passam a não ser suficientes informações sobre a situação interna atual da empresa, torna-se interessante incluir nas análises, dados sobre o ambiente externo, projeções, cenários, dados históricos. No sentido de melhor atender a estas necessidades, surgiram os Sistemas de Suporte a Decisão e Sistemas de Informação Executiva, descritos a seguir.

Sistemas de Suporte a Decisão e Sistemas de Informação Executiva

Outra categoria popular de sistemas de informação tem sido os sistemas de suporte a decisão - DSS e sistemas de informação executiva - EIS. Para diferenciá-los, define-se que DSS tendem a se focar mais em detalhes, sendo projetados para gerentes de baixo e médio escalão. EIS geralmente fornecem um maior nível de consolidação e uma visão multi-dimensional dos dados, dado que os executivos de alto nível necessitam de maior habilidade de dividir e agrupar os dados, do que funções para revisar os detalhes dos mesmos. (Gupta, 1997)

No entanto, decisões de nível estratégico e tático exigem um conteúdo contextual associado aos dados mais abrangente do que aquele encontrado nas bases operacionais. No nível estratégico, o executivo trabalha com decisões de longo prazo, geralmente um ou mais anos no futuro e, portanto, precisa ter uma boa visão dos fatores internos e externos que podem impactar sua empresa. No nível tático, gerentes de nível médio trabalham com decisões de médio prazo, normalmente de um mês a um ano. O trabalho destes gerentes consiste em amarrar as decisões estratégicas com as decisões operacionais. (Tecnologias, 1999)

Estas categorias de sistemas foram talvez os mais próximos precursores dos sistemas *data warehouse*, entretanto o alto custo do seu desenvolvimento e a coordenação necessária para sua produção acabaram por torná-los um produto de elite.

Gupta (1997) cita algumas características de DSS e EIS:

- ◊ tem seus dados em termos descritivos padrão de negócios, pois os nomes e estruturas nestes sistemas são projetados para uso por não técnicos;
- ◊ os dados geralmente são pré-processados com a aplicação de um padrão de negócios, tais como, distribuir os rendimentos por produtos, unidades de negócios e mercados;
- ◊ apresentam visões consolidadas dos dados, tais como produto, cliente e mercado. Ainda que estes sistemas possam ter a capacidade de descer aos dados mais detalhados, raramente são capazes de acessar todos os dados de detalhe ao mesmo tempo.

3.4 Data warehousing

Os sistemas *data warehouse* atuais fornecem todas as ferramentas analíticas oferecidas por seus precursores, mas seu projeto não é mais derivado de necessidades específicas de analistas ou executivos, obtendo maior sucesso quando seu projeto associa-se a toda a estrutura de negócios ao invés das necessidades específicas. (Gupta, 1997)

A idéia de *data warehouse* é integrar os dados internos e externos de uma organização em uma estrutura única, permitindo uma melhor utilização dos dados pelos analistas, gerentes e executivos. Uma vez obtida a integração, sistemas OLAP (*On Line Analytical Processing*) fornecem mecanismos sofisticados para análise dos dados.

Segundo Dantas e Filho (1997), quando se fala de *data warehouse* deve-se ter cuidado para não se fazer confusão. *Data warehouse* é apenas uma das modalidades de *data warehousing*, a mais robusta em termos de volumes de dados a serem armazenados. Portanto, tem-se, a seguir, as possíveis modalidades de soluções de *data warehouse* e suas características:

- *Data warehouse*:
 - ◊ Utilização altamente imprevisível, aplicações não-estruturadas, analíticas.
 - ◊ Tempo de resposta: segundos a alguns minutos.
 - ◊ Dados relacionais, não voláteis (*snapshots*), bastante desnormalizados.
 - ◊ Informações organizadas por área de análise, históricas (de cinco a dez anos).
 - ◊ Usuários finais: gerência, consumidores de informação.
- *Operational data store* (ODS):
 - ◊ Utilização previsível, parcialmente estruturada, parcialmente analítica.
 - ◊ Tempo de resposta: segundos a alguns minutos.
 - ◊ Dados relacionais, voláteis ou correntes, desnormalizados.
 - ◊ Informações organizadas por área de análise, históricas (de 30 a 60 dias).
 - ◊ Usuários finais: consumidores de informação.
- *Data mart*:
 - ◊ Tipo de *data warehouse* em que os dados estão mais próximos aos usuários.
 - ◊ Menores e mais fáceis de serem gerenciados.

- ◊ Permite tomada de decisões em nível departamental.
- ◊ Dados relacionais ou multidimensionais, não voláteis (*snapshots*).
- *OLAP servers*:
 - ◊ Menores que os *data warehouses*.
 - ◊ Bons recursos de exploração analítica.
 - ◊ Dados multidimensionais, não-voláteis (*snapshots*).
 - ◊ Solução complementar ao *data warehouse*.
- *Data mining*:
 - ◊ Baseado nos paradigmas de hipótese e descoberta.
 - ◊ Baseado em algoritmos de inteligência artificial.
 - ◊ Conceitos e ferramentas recentes.
 - ◊ Requer suporte intensivo do fornecedor.
 - ◊ Busca de padrão e comportamentos na base de dados.

Essas modalidades de implementação de *data warehousing* não são excludentes entre si. Cada organização escolherá aquela(s) que mais lhe convier(em) para obter as vantagens competitivas almejadas. Isso varia substancialmente entre os diversos segmentos de mercado, em diferentes graus de abertura e em diferentes graus de competitividade econômica de cada país ou região.

Segundo Gupta (1997), o desenvolvimento destas tecnologias está fortemente ligado ao desenvolvimento de outras tecnologias, tais como:

- ◊ *hardware* cada vez mais poderoso e a baixo custo;
- ◊ microcomputadores mais poderosos e com interfaces mais amigáveis;
- ◊ *softwares* dos equipamentos servidores cada vez mais estáveis, simples, poderosos;
- ◊ explosão de aplicativos para *internet* e *web*.

Ainda, segundo Gupta, (1997), um *data warehouse* pode ser caracterizado por:

- ◊ possuir dados armazenados em bancos de dados independentes dos sistemas operacionais;
- ◊ integrar dados de vários sistemas operacionais e *legacy systems*;
- ◊ possuir dados basicamente não voláteis (praticamente não sofrem alterações no decorrer do tempo);
- ◊ conter dados armazenados por longos períodos de tempo, para viabilizar análises temporais.

Em (Tecnologias, 1999) tem-se que um *data warehouse* é composto de um conjunto de dados :

- ◊ Orientados ou organizados por assunto. As bases operacionais são organizadas por aplicações e funções tais como, conta corrente, poupança, aplicação, etc. o *data warehouse* é organizado por assuntos tais como cliente, produto e fornecedor no caso de uma indústria, ou aluno, professor, e classe no caso de uma universidade.
- ◊ Integrados. Esta é a característica mais importante de um *data warehouse*. Numa empresa os dados estão em vários formatos (relacionais e não relacionais) e em diferentes locais, *mainframe*, sistemas cliente-servidor, servidores departamentais, fornecedores externos, enfim, em toda parte. Além da integração de todos estes dados espalhados, os dados no *data warehouse* são submetidos a um processo de limpeza, condicionamento e reconstituição de chaves. A tabela a seguir mostra exemplos da falta de integração, consistência e integridade dos dados em um ambiente operacional. Como pode-se observar, as diversas aplicações utilizam unidades, códigos e valores diversos para a representação dos mesmos tipos de dados. Para a consolidação das informações no *data warehouse*, é necessária a definição de um padrão e a adequação dos valores encontrados nas bases operacionais, o que é feito durante o processo de carga dos dados para o *data warehouse*.

Operacional	Data warehouse
Aplic A - altura cm Aplic B - altura m Aplic C - altura poleg	altura cm
Aplic A - m,f Aplic B - 0,1 Aplic C - h,m Aplic D – masculino,feminino	m,f
Aplic A - saldo Aplic B - saldo_disp Aplic C - sal_disp Aplic D - disponivel	saldo

Tabela 1: falta de integração, consistência e integridade dos dados em um sistema operacional.

- Não-voláteis. No ambiente operacional os dados mudam a todo instante, pois operações de *insert*, *update* e *delete* são feitas constantemente na base. Para a tarefa de análise, é essencial que os dados tenham menos volatilidade e por isto, em um *data warehouse*, as operações básicas são a de acesso aos dados e de carregamento (*load*) da base, sendo esta última, executada, por exemplo, uma vez ao dia e em um horário quando a base não é acessada por nenhum usuário.
- Variáveis no tempo (ou históricos). Enquanto no ambiente operacional é armazenado apenas o valor mais recente do dado, em um *data warehouse* são mantidos todos os valores que o dado teve dentro de um horizonte de tempo específico. Dados históricos são importantes para identificar tendências.

Além destas características, o *data warehouse* diferencia-se dos sistemas OLTP também em outros aspectos. Enquanto nos sistemas OLTP existem dezenas de usuários acessando os dados, no *data warehouse* o número de usuários é bem menor. Nos sistemas OLTP as consultas são pré-definidas. No *data warehouse*, elas são complexas, imprevisíveis e geralmente envolvem um grande volume de dados. Apesar disso, o tempo de resposta não é tão crítico como nos sistemas OLTP

Qual seria a modalidade de *data warehouse* mais apropriada às necessidades e objetivos competitivos da organização? Qual a melhor abordagem de implantação a ser adotada? Quais devem ser os requisitos para a escolha de soluções de *data warehousing*? Essas são, e provavelmente serão, as perguntas mais

freqüentemente formuladas entre os tomadores de decisão interessados numa implementação de *data warehousing*. A fim de nortear as respostas a tais perguntas, estão listadas a seguir alguns critérios: (Dantas e Filho, 1997)

- ◊ *Data warehouses* são soluções geralmente adotadas por grandes corporações, cuja combinação das variáveis competitividade e volume de negócios e informações justifiquem o porte desse tipo de solução. Estudos feitos entre os elencados na Fortune 500 apontam que 95% dessas empresas já implementaram ou estão pensando em implementar seus *data warehouses*. Esse tipo de implementação, tipicamente, envolve um horizonte de dois ou três anos e investimentos da ordem de seis a sete milhões de dólares.
- ◊ *Data marts* são soluções apropriadas às grandes e médias corporações, dado que representam um tipo de *data warehouse* menos complexo em termos de implementação e mais simples de ser gerenciado, pois tem requisitos menos complexos em termos de infra-estrutura e de abrangência funcional.
- ◊ *OLAP servers* são soluções interessantes como ponto de partida numa implantação mais abrangente de *data warehouses* ou *data marts*, além de terem o porte sob medida às necessidades das médias e pequenas organizações, incluindo aquelas em segmentos de negócio com menor pressão competitiva.
- ◊ Implementações de *data warehouses* devem ser incrementais, a fim de possibilitarem a obtenção de resultados em prazos menores (três a quatro meses) e a justificativa dos investimentos necessários. A abordagem deve ser iniciada com etapas preliminares de definição funcional do escopo do projeto e de prototipagem que demonstrem a viabilidade técnica das arquiteturas propostas.

Como citado anteriormente, são vários os fatores que levam as empresas a optarem por soluções de implantação gradual, com custos menores e resultados visíveis a mais curto prazo. É importante que os primeiros resultados sejam disponibilizados a curto prazo. Com este objetivo, deve-se traduzir rapidamente as necessidades do negócio em uma especificação que possa ser construída em etapas, para minimizar os riscos e o tempo de apresentação dos resultados iniciais.(Unicamp, 1996b) Neste contexto cresce o interesse e o uso de

data marts, dentre as modalidades de *data warehousing*. Este processo, bem como formas de implementação e motivos para esta opção serão detalhados no item que segue.

3.5 Data marts

Os primeiros projetos *data warehouse* implementavam uma arquitetura centralizada (figura 1). Esta arquitetura assegurava uniformidade, controle e maior segurança, entretanto a implementação desta abordagem não é uma tarefa fácil, nem rápida ou de baixo custo. Exige uma metodologia rigorosa e uma completa compreensão dos negócios da empresa.

Com o aparecimento de *data mart* ou *warehouse* departamental, surgiu a abordagem descentralizada (figura 2) como uma das opções de arquitetura *data warehouse*. Com o uso desta arquitetura podem ser obtidos, custo mais baixo e implementação mais rápida. As desvantagens estão no maior número de ações de extração e ou transformação dos dados das bases operacionais para os *data marts* e também no risco da proliferação de *data marts* sem controle, gerando problemas de integração. (Tecnologias, 1999)

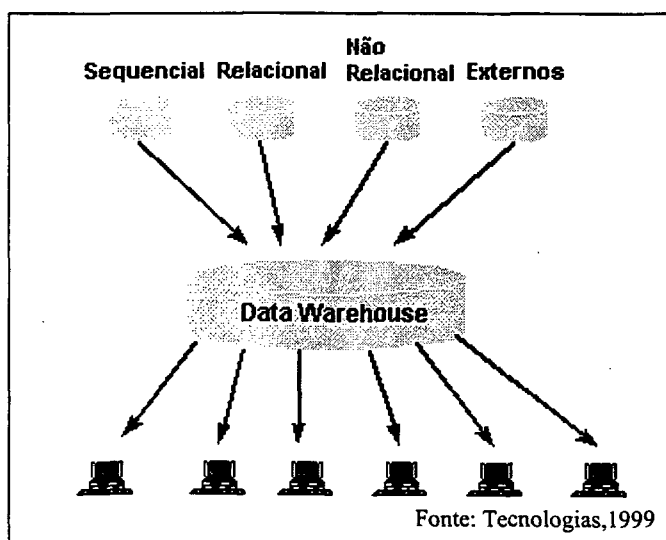


Figura 1 - Arquitetura Centralizada

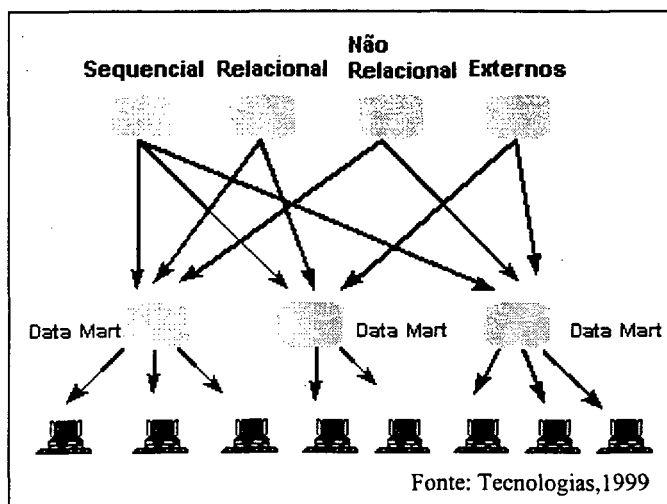


Figura 2 - Arquitetura descentralizada

Segundo Jones (1998), “Na virada do século, os *data marts* irão dominar o mundo do suporte à decisão. Conforme algumas estimativas, os *data marts* irão contabilizar três quartos de todos os sistemas de BI (*business intelligence*) feitos nos próximos anos.” No mesmo artigo, Jones (1998) afirma que isto se dará por serem os *data marts* menores, de custo menor e mais direcionados (*focused*) que *data warehouses* de larga escala. O autor afirma que somente as vantagens econômicas são suficientes para garantir sua popularidade.

Os *data marts* foram propostos no início da década de 90 como estruturas complementares aos *data warehouses* centralizados. Originalmente, o *data warehouse* coletaria e unificaria todas as informações para o suporte a decisão da organização, servindo os *data marts* como mecanismos para acesso dos usuários finais. Entretanto, na medida em que os projetos de *data warehouse* são muito caros, demorados e com maior risco para muitas organizações, o *data mart* foi sendo redefinido. Atualmente, o *data mart* é visto como uma alternativa viável ao *data warehouse* centralizado. Esta mudança no significado de *data mart* tem causado percepções equivocadas e muita confusão. O engano mais comum sobre *data marts* é que eles não podem suportar uma visão corporativa da informação. Sem dúvida, algumas empresas têm tido problemas na implantação da estratégia de *data mart*,

geralmente quando as unidades de negócios (BUs) constroem *data marts* autônomos, sem considerar integração, mas é uma grande simplificação dizer que por que algumas empresas não dedicaram o tempo necessário para integrar seus *data marts*, a tarefa é impossível.

Na realidade, *data marts* podem ser desenvolvidos e integrados gradualmente, sem que um *data warehouse* centralizado seja feito primeiro. A chave para esta estratégia é a aplicação adequada da modelagem dimensional e de princípios do projeto. Mais especificamente, deve-se assegurar que os dados com dimensão organizacional, tais como clientes e produto, sejam uniformes em todos os *data marts*. Desta forma, quando for necessário ou desejado, podem ser feitas consultas integrando um ou mais *data marts*, sendo que os resultados serão consistentes.

Não é completamente sem sentido argumentar que uma estratégia de *data mart* efetiva não é nada mais que construir um *data warehouse* de forma incremental. Entretanto, argumentações sobre terminologia são perda de tempo, os ambientes de negócios atuais demandam por uma capacidade de suporte a decisão rápida e a custos aceitáveis (*cost-effective*), estando esta disponível, não importa como tal tecnologia é chamada.

3.6 Definição de um modelo de *data warehouse*

"Um *data warehouse* não é um produto que se compra, mas sim projeto que envolve análise e implementação, com a participação de várias tecnologias." (Unicamp, 1996b) Neste sentido, encontram-se várias propostas de metodologia e formas de implementação de *data warehouses*.

O processo de modelagem de dados para a estruturação de um *data warehouse* deve ser independente do modelo de dados relacional utilizado nos sistemas operacionais. O modelo do *data warehouse* tende a ser bem menos

normalizado que o dos sistemas operacionais, bem como atributos¹ essenciais aos sistemas operacionais podem ser desnecessários no *data warehouse*. Da mesma forma o modelo do *data warehouse* deve ser extensível e estruturado de modo que dados de diferentes aplicações possam ser acrescentados. A maioria dos projetos de sucesso, utilizam-se de uma abordagem incremental para adicionar dados dos sistemas operacionais. Os projetos são iniciados com o objetivo de acrescentar quase todos, se não todos os dados do negócio. Mantendo este projeto em mente, pode-se começar com uma ou duas aplicações operacionais que forneçam os dados mais interessantes para as análises do negócio. (Gupta, 1997).

De forma bastante resumida, pode ser dito que os principais passos para a definição de um *data warehouse* são: 1) determinar os dados que popularão o *data warehouse* e quais as fontes de dados que serão necessárias; 2) armazenar e gerenciar estes dados, garantindo confiabilidade e acesso adequados; e 3) a disponibilização destes dados para o suporte à decisão.

Todo este processo de análise e definição do modelo deve estar vinculado aos processos do negócio, através de um processo simples e compreensível de definição do *data warehouse*. Partindo dos processos do negócio, o analista tende a se obrigar a produzir um modelo mais voltado às reais necessidades da empresa, e não a um problema ou demanda específicos da ocasião do procedimento de análise e projeto. Neste sentido, são apresentados a seguir dois conjuntos de passos, o primeiro proposto por Kimball (1998) e o segundo por Miley (1998) como orientação a ser seguida na definição de um *data warehouse*.

Em Kimball (1998), encontram-se nove pontos de decisão para o projeto de um *data warehouse*, propondo cada um deles a identificação de:

- Os processos, que vão definir a identidade das tabelas de fatos. Como citado anteriormente, a modelagem dimensional do *data warehouse* é um processo de projeto *top down*, em que partimos da definição do negócio para escolher quais as dimensões e fatos devem constar do banco de dados. Um processo é uma

¹ detalhes utilizados para qualificar, identificar, classificar ou quantificar um objeto ou entidade.

operação importante da organização, normalmente suportada por algum tipo de sistema legado, de onde será possível extrair dados para o *data warehouse*. A seleção do processo a ser modelado deve ser feita a partir de informações sobre o negócio e os dados disponíveis.

- ◊ A granularidade, ou o nível de detalhe de cada tabela de fatos. O maior nível de detalhe do fato a ser armazenado no *data warehouse*, corresponde ao grão. Grãos típicos correspondem a transações, instantâneos diários ou mensais.
- ◊ As dimensões de cada tabela de fatos. As dimensões correspondem aos elementos que caracterizam a ocorrência de um fato. Por exemplo, em um *data warehouse* que modele fatos de vendas, tendo como grão a movimentação diária de um item, poderia ter como dimensões: tempo, produto e loja. Neste caso, uma ocorrência na tabela de fatos associaria a quantidade vendida de um produto em uma loja em uma determinada data.
- ◊ Os fatos. A tabela de fatos armazena medições do negócio, sendo caracterizada pela intersecção de todas as dimensões. Mantendo o exemplo do *data warehouse* de vendas, os fatos poderiam corresponder ao valor total da venda, número de unidades vendidas e custo estendido dos itens vendidos (custo unitario x quantidade).
- ◊ os atributos de cada dimensão, com descrições completas e terminologia apropriada. Os atributos destinam-se a descrever os itens de uma dimensão, sendo utilizados como restrições e cabeçalhos de linha nos resultados das consultas ao *data warehouse*. Por exemplo, atributos típicos de um produto poderiam ser descrição, marca, categoria, tipo de embalagem e tamanho.
- ◊ Forma de tratamento das dimensões de modificação lenta. As dimensões, como produto e cliente, não são independentes em relação ao tempo: a formulação dos produtos evolui; as pessoas modificam seus nomes, casam e se divorciam, aumentam o número de filhos, mudam de endereço. Estas modificações podem ser tratadas de três formas: 1) substituindo os valores antigos pelos novos, perdendo a capacidade de rastrear o histórico passado; 2) adicionando um registro à dimensão contendo os novos valores no momento da mudança, segmentando o histórico entre a descrição antiga e a nova; e 3) criando novos campos para os

valores atuais, no registro original, para a inclusão dos novos valores, mantendo também os valores originais, desta forma são mantidos o histórico anterior e posterior à mudança. Cada uma destas abordagens tem seus prós e contras, devendo cada caso ser analisado e feita a opção mais conveniente.

- ◊ Os agregados, dimensões heterogêneas, minidimensões, modos de consulta e outras decisões de armazenamento físico. Muitas consultas e relatórios feitos sobre o *data warehouse* são operações simples, baseadas em parâmetros predefinidos. Como exemplo, em uma organização, muitas pessoas podem precisar de informações sobre vendas de produtos. Eles podem ter a necessidade de totalizar dados por semana ou mês, e pode não ser prático calcular os totais a cada consulta. Um *data warehouse* que contenha visões totalizadas dos dados das consultas mais freqüentes pode reduzir grandemente o processamento necessário no momento das análises.
- ◊ A amplitude de tempo a ser mantida no histórico do banco de dados. Nos sistemas operacionais, os dados são arquivados após tornarem-se inativos, por exemplo, os dados sobre uma conta bancária podem ser excluídos após um período de inatividade da conta. A razão principal para o arquivamento dos dados inativos tem sido a performance dos sistemas operacionais. No caso dos *data warehouses*, o projeto pode iniciar sem um plano de arquivamento de dados. O custo de manutenção dos dados após sua carga no *data warehouse* é mínimo. Entretanto, é importante fazer uma estimativa do tamanho do *data warehouse* para o dimensionamento do *hardware* e *software* necessários.
- ◊ Os intervalos de extração de dados e carga do *data warehouse*. Da definição destes intervalos depende a atualização dos dados que serão disponibilizados. Para esta decisão devem ser ponderados o impacto da ausência de dados mais recentes nas bases de consulta e o custo de intervalos de extração e carga menores.

Em Miley (1998), encontram-se uma lista de itens importantes no desenvolvimento de uma solução de *data warehouse*, qual seja:

- ◊ Determinação das necessidades dos usuários finais e construção de um modelo das questões de negócio a serem respondidas;
- ◊ Identificação dos dados de fontes internas e externas que irão alimentar o *data warehouse*;
- ◊ Análise dos dados e das fontes para construir um modelo de funções e processos envolvidos, com o objetivo de determinar as regras que regem o negócio e definir o nível dos elementos de dados necessários no *data warehouse*;
- ◊ Determinação dos processos e da integração entre os dados definidos nas fontes e os definidos no *data warehouse*, detectando o volume de dados a extrair e transformar, a periodicidade de atualização e os eventos que devem disparar uma atualização;
- ◊ Definição dos metadados, que definem as fontes, transformações e a lógica para o *data warehouse* - metadados são a descrição do que está no banco de dados e como ele está armazenado, sendo que um repositório de metadados deve conter as definições dos dados, regras de negócio e a lógica detalhada para modelagem do desenvolvimento dos sistemas;
- ◊ Construção e população do *data warehouse*;
- ◊ Construção dos *data marts* que irão conter os dados sumarizados e subconjuntos do *data warehouse*, sendo que parte dos metadados irá definir como dados primários no *data warehouse* serão transformados, sumarizados e armazenados nos *data marts*;
- ◊ Disponibilização da aplicação OLAP e dos *web servers* com o que for necessário para os acessos, análises, relatórios e transações dos usuários;
- ◊ Instalação dos *web browsers* e *plug ins* necessários nos PCs dos usuários.

3.7 Ferramentas para recuperação e análise de dados

O valor de um *data warehouse* está na qualidade das decisões que ele suporta, e as ferramentas de análise de informação utilizadas são importantes para aumentar o desempenho destes sistemas. As três abordagens básicas que podem ser empregadas, serão descritas nos itens que seguem: consultas convencionais, OLAP, e *data mining* (Unicamp, 1996b).

3.7.1 Consultas convencionais

São consultas e relatórios realizados através de ferramentas genéricas ou de implementação própria, não desenvolvidas com o objetivo ou abordagem próprios para consultas a *data warehouses*. Estas ferramentas geralmente permitem ao usuário a execução de consultas simples adequando-se a sistemas de decisão de pequeno porte. Suas principais limitações são:

- ◊ dificuldade com consultas complexas, tanto em sua elaboração pelo usuário, quanto na otimização da sua execução;
- ◊ ausência dos conceitos de tempo, consolidação e agregação;
- ◊ necessidade e apoio de pessoal técnico para elaboração de consultas pela necessidade de conhecimento do projeto da base de dados; e
- ◊ em muitas destas ferramentas a manipulação dos dados é executada na máquina do usuário, gerando sobrecarga tanto na estação quanto a rede.

3.7.2 OLAP

As ferramentas OLAP (*on line analytical processing*) utilizam-se de estruturas multidimensionais, ou seja, mostram os dados no formato de uma planilha e permitem sua visualização em diferentes perspectivas. Estas ferramentas foram projetadas para suportar análise e consultas *ad hoc*², ajudando analistas e executivos a sintetizarem informações sobre a empresa através de comparações, visões, análise histórica e projeção de dados com elaboração de cenários. Principais características:

- ◊ Visão dimensional dos dados;
- ◊ Análise de tendências e elaboração de cenários;
- ◊ Busca automática de detalhes;
- ◊ Dimensionalidade genérica;
- ◊ Operação transdimensional, ou seja, possibilidade de fazer cálculos e manipulação de dados através diferentes dimensões;

² consultas *on line* não planejadas (casuais).

- ◊ Diversas possibilidades de visualização dos dados através da mudança de ponto de vista e navegação entre os níveis de agregação (*drill-up/drill-down*³); e
- ◊ Conjunto de funções de análise e cálculos não triviais com os dados.

3.11.3 Data mining

Ferramentas de *data mining* são aquelas que fazem a exploração de um grande volume de dados, procurando padrões, tendências e relacionamentos entre os dados. *Data mining* não é uma tecnologia nova, no entanto tomou impulso com o aparecimento do *data warehouse*.

A diferença básica entre ferramentas OLAP e *data mining* está na maneira de exploração dos dados. Nas ferramentas OLAP, o analista conhece a questão, elabora uma hipótese e utiliza a ferramenta para rejeita-la ou confirmá-la. Com *data mining*, a ferramenta é utilizada para a busca de conhecimento. Esta busca é feita buscando-se identificar:

- ◊ Associações entre os dados, procurando identificar afinidades ou padrões entre um conjunto de dados. Por exemplo, quais produtos são comprados em conjunto?;
- ◊ Grupos (*Clusters*), consiste na procura de grupos de dados de acordo com algum critério. Exemplo, quais produtos têm mais defeitos ?;
- ◊ Classificações, busca a identificação de regras que descrevem a qual grupo um item pertence, são examinados itens já classificados e inferido um conjunto de regras. Por exemplo, se do sexo masculino, idade entre 18 e 30 e carro esportivo, então alta chance de ter um acidente; e
- ◊ Padrões sequenciais, procura tendências na análise de um conjunto de dados em um determinado intervalo de tempo. Exemplo, se evento A ocorre, então X% das vezes, evento B ocorre.

³ *Drill down* significa mostrar mais detalhes, o que em um *data warehouse* corresponde a adicionar cabeçalhos de linha de tabelas de dimensão. *Drill up* é subtrair cabeçalhos de linha. (Kimball, 1998).

3.12 *Data warehouse e Web*

O assunto *data warehouse* tem sido constantemente tratado na literatura devido a sua importância estratégica dentro da empresa, por fornecer dados confiáveis e de fácil acesso. Neste sentido, o processo de construção de um *data warehouse* passa incluir a tecnologias de *internet/intranet*, vislumbrando a disponibilização dos dados através de navegadores (*browsers*) acoplados a ferramentas de análise de dados (Unicamp, 1996b).

Data marts e *data warehouses* vêm sendo considerados por alguns como a nova geração de DSSs. E na medida em que este movimento vai se realizando, muitos voltam-se a *internet* para disponibilizar esta inteligência de negócios a uma audiência mais ampla de trabalhadores do conhecimento. O movimento no sentido de um modelo de negócios cliente-cêntrico e seus vários envolvidos, em um ambiente de competição global crescente tem acarretado o desenvolvimento de aplicações DSS para os trabalhadores de linha de frente, tais como o pessoal de vendas no varejo e os serviços de clientes, que gerenciam as carteiras de clientes. (Miley, 1997)

Segundo Miley (1997), as razões para disponibilizar o *data warehouse* e *data mart* na *internet/intranet* seriam:

- ◊ baixo custo de desenvolvimento e gerenciamento, na medida em que minimiza a necessidade de instalação e manutenção nas estações clientes;
- ◊ acesso universal a informação e aplicações, de forma independente de plataforma ou localização - a informação torna-se disponível em qualquer máquina conectada a *internet* e que possua um *web browser*;
- ◊ sistemas de informação *self service*, nos quais os usuários podem acessar a informação de que precisam quando precisam, independente da necessidade de um analista ou programador para preparar o relatório;
- ◊ redução dos custos de treinamento, dado que a partir do momento em que aprender a utilizar o *web browser*, o usuário está capacitado para acessar e usar qualquer dos recursos da *intranet*, *extranet* ou *internet*.

A distribuição de informações na *www*⁴ proporciona ganhos significativos para a empresa na facilidade de acesso (de qualquer lugar) e na maior rapidez de atualização e distribuição da informação, sendo que os EIS também estão sendo disponibilizados na *web*.

Ainda segundo Miley (1997), o *data warehouse* na *web* tem três domínios primários:

- ◊ *intranet*: para suporte a decisão dentro da empresa, sendo que qualquer empresa que possua escritórios geograficamente dispersos tende a beneficiar-se de disponibilizar seu *data warehouse* na *web*. Por ex., um banco pode disponibilizar um EIS baseado no *data warehouse* para o planejamento financeiro estratégico, fornecendo ferramentas OLAP sofisticadas para os analistas executivos efetuarem análises de dados multidimensionais; e para os representantes junto a clientes fornecer relatórios e consultas padrão disponíveis através do *web browser*.
- ◊ *extranet*: para suporte a clientes, bem como a parceiros acima e abaixo da cadeia de fornecimento. Qualquer empresa de serviços ou largamente envolvida em transações como distribuidores, agências de viagem, autolocadoras, tende a beneficiar-se de uma *extranet*. Por exemplo, um fornecedor de vestuário ligado por uma rede WAN a um vendedor franqueado, poderia utilizar um *web browser* para acessar os pedidos e dados de estoque consolidados por região, neste caso o fornecedor poderia utilizar esta informação para prever novos fornecimentos de determinados produtos para uma região de atendimento em particular.
- ◊ *internet*: para *marketing* de produtos ou *data warehouses* geradores de lucros, onde partes do DSS são vendidas a grupos selecionados mediante o pagamento de taxas.

Assim, cada vez mais DSS deixa de ser somente OLAP, passa a incluir a *web* e, eventualmente, *groupware*, *workflow* e novas fontes para um suporte mais completo a decisão.

⁴ World wide web

4. METODOLOGIA

Para atingir os objetivos do trabalho foi definido o plano de investigação onde é determinada a forma como será feita a coleta de dados e a análise das informações. Inicialmente, é feita uma revisão bibliográfica para estabelecer uma visão do assunto, seguindo-se uma parte teórica onde estes conhecimentos são aplicados.

4.1 Delineamento de pesquisa

Este trabalho tem uma abordagem qualitativa. Quanto aos fins é uma pesquisa exploratória aplicada. Trata-se de uma pesquisa exploratória, uma vez que diz respeito a tecnologias novas, sobre as quais ainda há pouco conhecimento acumulado e sistematizado, buscando-se esclarecer e contextualizar o uso destes novos conceitos e ferramentas.

A motivação desta pesquisa é a necessidade do domínio das tecnologias de *data warehousing*. Sendo desenvolvida uma parte prática onde estes conhecimentos são empregados, contextualizando e mostrando a aplicabilidade destas novas tecnologias na tentativa de solução de um problema prático do SEINFO do TRT.

Quanto aos meios, este projeto é ao mesmo tempo uma pesquisa bibliográfica e estudo de caso. Inicialmente são estudadas as tecnologias de *data warehousing*, com base em material publicado em livros, revistas e na *internet*, e a partir das informações adquiridas, é desenvolvido o modelo proposto neste trabalho.

4.2 Delimitação da população pesquisada

O universo da pesquisa corresponde ao conjunto de pessoas que possuem informações sobre os processos desenvolvidos na área de informática. Neste grupo, estão diretores, chefes, e os demais servidores da referida área. Deste universo, uma população amostral foi escolhida para fornecer as informações necessárias ao desenvolvimento do modelo. A seleção dos elementos desta amostra foi feita sem a utilização de métodos estatísticos. Foram selecionados técnicos dos setores de administração de bancos de dados e de administração de sistemas operacionais, sendo as pessoas escolhidas por acessibilidade, sendo entrevistados os elementos aos quais o acesso era mais conveniente.

4.3 Técnicas de coleta de dados

Foram coletados dados primários através de observação participante e de entrevistas não estruturadas. A observação participante tornou-se conveniente pelo fato de a autora fazer parte do quadro de informática do TRT. As entrevistas foram feitas discutindo-se com o entrevistado sobre o problema a ser investigado, de modo a obter informações sobre ele. Na pesquisa bibliográfica, são obtidos dados secundários, a partir de livros, revistas e da *internet*, sobre as tecnologias de *data warehousing*.

4.4 Tratamento dos dados

Os dados obtidos são tratados de forma qualitativa. As informações obtidas, através do levantamento bibliográfico, da observação e das entrevistas, foram organizadas de forma a descrever o ambiente do tema-problema diante das novas tecnologias existentes e da situação da informatização do TRT. A organização destes visa atender adequadamente aos propósitos do projeto, permitindo a análise adequada dos dados.

4.5 Limitações do método

A população amostral selecionada para as entrevistas não contempla todos os possíveis usuários do sistema, de forma que após a implantação do modelo pode surgir a necessidade de adequação do mesmo a necessidades não previstas. Mas, como a implantação de um *data warehouse* é um processo, que não termina, requerendo contínuo monitoramento e validação, a manutenção e adaptação do mesmo fazem parte do sistema.

Atualmente, não existe nenhum modelo de *data warehousing* ou qualquer outro sistema de propósitos similares ao proposto, desenvolvido ou em desenvolvimento no TRT, de forma que não há experiência acumulada no assunto.

As tecnologias estão em evolução constante, de forma que, quando da implantação do modelo, possivelmente existirão outras alternativas que satisfaçam o problema.

A biblioteca da UFSC possui pouca bibliografia sobre o tema, sendo que a maior parte do material encontrado são artigos de revistas ou artigos encontrados na *internet*.

5. APRESENTAÇÃO E ANÁLISE DOS DADOS

Nesta parte do trabalho, será feita uma caracterização do TRT, seguida da apresentação do desenvolvimento do modelo de *data mart* proposto.

5.1 Caracterização da organização

Na Constituição de 1934, foi encontrada uma das primeiras definições da Justiça do Trabalho, como se observa em seu conteúdo, apresentado a seguir:

Seção VI

Dos Juizes e Tribunais do Trabalho

Art.122 - Os órgãos da Justiça do Trabalho são os seguintes:

I - Tribunal Superior do Trabalho;

II - Tribunais Regionais do Trabalho;

III - Juntas ou Juizes de Conciliação e Julgamento.

§ 1º - O Tribunal Superior do Trabalho tem sede na Capital federal.

§ 2º - A lei fixará o número dos Tribunais Regionais do Trabalho e respectivas sedes.

§ 3º - A lei instituirá as Juntas de Conciliação e Julgamento podendo, nas Comarcas onde elas não forem instituídas, atribuir as suas funções

aos Juizes de Direito.

§ 4º - Poderão ser criados por lei outros órgãos da Justiça do Trabalho.

§ 5º - A constituição, investidura, jurisdição, competência, garantias e condições de exercício dos órgãos da Justiça do Trabalho serão reguladas por lei, ficando assegurada a paridade de representação de empregados e empregadores.

Art. 123 - Compete à Justiça do Trabalho conciliar e julgar os dissídios individuais e coletivos entre empregados e empregadores, e as demais controvérsias oriundas de relações do trabalho regidas por legislação especial.

§ 1º - Os dissídios relativos a acidentes do trabalho são da competência da Justiça ordinária.

§ 2º - A lei especificará os casos em que as decisões, nos dissídios coletivos, poderão estabelecer normas e condições de trabalho."

Em sua estrutura inicial, a Justiça do Trabalho dividia-se nas mesmas três instâncias de hoje: as Juntas de Conciliação e Julgamento, os Conselhos Regionais do Trabalho e o Conselho Nacional do Trabalho. A Constituição de 1946 integrou-a ao Poder Judiciário, como ramo especializado. Os Conselhos Regionais tornaram-se Tribunais Regionais e o Conselho Nacional passou a ser o Tribunal Superior do Trabalho.

Já, na Constituição Federal de 1988, encontra-se a definição da competência da Justiça do Trabalho, como segue:

Compete à Justiça do Trabalho conciliar e julgar os dissídios individuais e coletivos entre trabalhadores e empregadores, abrangidos os entes de direito público externo e da administração pública direta e indireta dos Municípios, do Distrito Federal, dos Estados e da União, e, na forma da lei, outras controvérsias decorrentes da relação de trabalho, bem como os litígios que tenham origem no cumprimento de suas próprias sentenças, inclusive coletivas. (art. 114).

Desta forma, tudo o que se relacionar com conflitos entre trabalhadores e empregadores, no plano individual ou coletivo, resultante da relação de emprego, é da competência da Justiça do Trabalho. Cabe, assim, a seus órgãos, nos diferentes graus de jurisdição, decidir sobre tais conflitos. Os servidores públicos, porém, por decisão do Supremo Tribunal Federal, estão

fora da competência da Justiça do Trabalho (ADIN nº492, de 12.11.92).

Para a solução de dissídio coletivo (de interesse de categorias profissionais ou econômicas), a Justiça do Trabalho pode "estabelecer normas e condições, respeitadas as disposições convencionais e legais mínimas de proteção ao trabalho" (art.114, §2º). É o chamado poder normativo.

Ainda na página *internet* do TST, encontra-se uma descrição bastante clara da estrutura da Justiça do Trabalho, como segue:

Desde a sua criação, a estrutura e a organização permaneceram praticamente inalteradas. Existem na Justiça do Trabalho três graus de jurisdição:

Primeiro Grau - Juntas de Conciliação e Julgamento (JCJs): Julgam apenas dissídios individuais. Sua jurisdição é local (abrange geralmente um ou alguns municípios). Em comarcas onde não exista JCJ, a função é exercida pelo Juiz de Direito. Existem atualmente 1.109 JCJs no País.

A JCJ compõe-se de um Juiz Presidente (juiz de carreira) e dois juizes classistas, um indicado pelos empregados, outro pelos empregadores.

Dissídios individuais são controvérsias surgidas nas relações de trabalho entre o empregador (pessoa física ou jurídica) e o empregado (este sempre como indivíduo, pessoa física). Esse conflito chega à JCJ na forma de Reclamação Trabalhista.

Segundo Grau - Tribunais Regionais do Trabalho (TRTs): Julgam recursos ordinários contra decisões de Juntas, agravos de instrumento, ações originárias (dissídios coletivos de categorias de sua área de jurisdição - sindicatos patronais ou de trabalhadores organizados em nível regional), mandados de segurança, ações rescisórias de decisões suas ou das JCJs etc.

A Constituição Federal de 1988 estabelece que "haverá pelo menos um TRT em cada Estado e no Distrito Federal" (art.112).

Atualmente, existem 24 TRTs. Em São Paulo existem dois, um na Capital, outro em Campinas. Não foram ainda criados TRTs nos Estados de Tocantins, Acre, Roraima e Amapá.

Os juízes dos TRTs são nomeados pelo Presidente da República e seu número varia em função do volume de processos examinados pelo Tribunal, mantida, porém, a proporção de dois terços de juízes togados (vitalícios) e um terço de classistas (temporários), com representação paritária de empregados e empregadores. Nenhum TRT tem menos de oito juízes.

Terceiro Grau - Tribunal Superior do Trabalho (TST): O TST tem por principal função uniformizar a jurisprudência trabalhista. Julga recursos de revista, recursos ordinários e agravos de instrumento contra decisões de TRTs e dissídios coletivos de categorias organizadas em nível nacional, além de mandados de segurança, embargos opostos a suas decisões e ações rescisórias. É composto por 27 ministros, nomeados pelo Presidente da República, sendo 17 togados (vitalícios) e dez classistas (temporários), cinco representantes de empregados e cinco de empregadores.

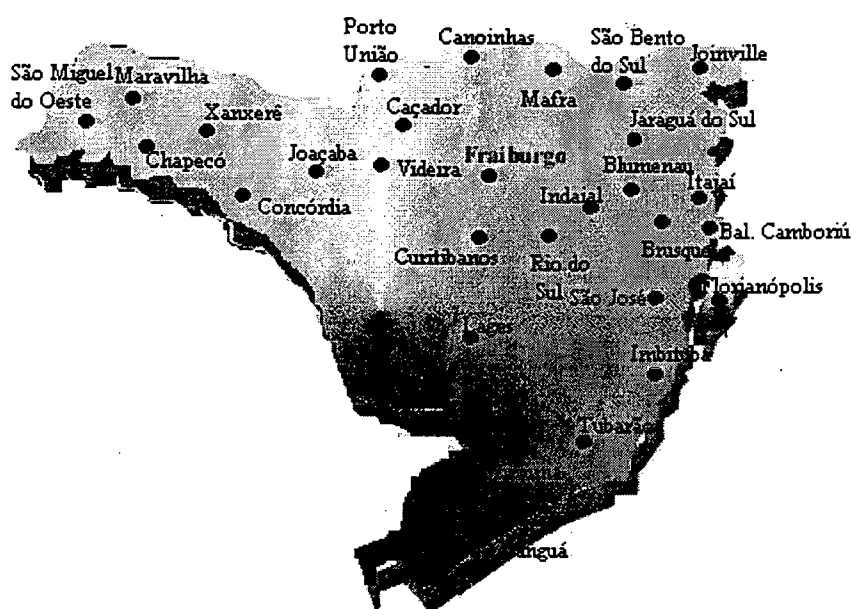
Quanto às Unidades Judiciárias do TRT - 12ª Região, podem-se citar:

- ◊ JCJ de Araranguá
- ◊ JCJ de Balneário Camboriú
- ◊ Serviço de Distribuição dos Feitos de 1ª Instância de Blumenau
- ◊ 1ª JCJ de Blumenau
- ◊ 2ª JCJ de Blumenau
- ◊ 3ª JCJ de Blumenau
- ◊ JCJ de Brusque
- ◊ JCJ de Caçador
- ◊ JCJ de Canoinhas
- ◊ Serviço de Distribuição dos Feitos de 1ª Instância de Chapecó
- ◊ 1ª JCJ de Chapecó
- ◊ 2ª JCJ de Chapecó
- ◊ JCJ de Concórdia
- ◊ Serviço de Distribuição dos Feitos de 1ª Instância de Criciúma
- ◊ 1ª JCJ de Criciúma
- ◊ 2ª JCJ de Criciúma
- ◊ JCJ de Curitiba

- ◊ Serviço de Distribuição dos Feitos de 1ª Instância de Florianópolis
- ◊ 1ª JCJ de Florianópolis
- ◊ 2ª JCJ de Florianópolis
- ◊ 3ª JCJ de Florianópolis
- ◊ 4ª JCJ de Florianópolis
- ◊ 5ª JCJ de Florianópolis
- ◊ 6ª JCJ de Florianópolis
- ◊ 7ª JCJ de Florianópolis
- ◊ JCJ de Imbituba
- ◊ JCJ de Indaial
- ◊ Serviço de Distribuição dos Feitos de 1ª Instância de Itajaí
- ◊ 1ª JCJ de Itajaí
- ◊ 2ª JCJ de Itajaí
- ◊ JCJ de Jaraguá do Sul
- ◊ JCJ de Joaçaba
- ◊ Serviço de Distribuição dos Feitos de 1ª Instância de Joinville
- ◊ 1ª JCJ de Joinville
- ◊ 2ª JCJ de Joinville
- ◊ 3ª JCJ de Joinville
- ◊ 4ª JCJ de Joinville
- ◊ Serviço de Distribuição dos Feitos de 1ª Instância de Lages
- ◊ 1ª JCJ de Lages
- ◊ 2ª JCJ de Lages
- ◊ JCJ de Mafra
- ◊ JCJ de Porto União
- ◊ JCJ de Rio do Sul
- ◊ JCJ de São Bento do Sul
- ◊ Serviço de Distribuição dos Feitos de 1ª Instância de São José
- ◊ 1ª JCJ de São José
- ◊ 2ª JCJ de São José
- ◊ JCJ de São Miguel do Oeste

- ◊ UJA - Unidade Judiciária Avançada de Maravilha
- ◊ Serviço de Distribuição dos Feitos de 1ª Instância de Tubarão
- ◊ 1ª JCJ de Tubarão
- ◊ 2ª JCJ de Tubarão
- ◊ JCJ de Videira
- ◊ UJA - Unidade Judiciária Avançada de Fraiburgo
- ◊ JCJ de Xanxerê

O mapa de Santa Catarina, a seguir, mostra como as unidades judiciárias do TRT/SC estão distribuídas ao longo do estado:



No entanto, vale lembrar que, neste trabalho, o alvo é somente a área de informática do Tribunal, compreendendo a Assessoria de Informática (ASINFO) e o Serviço de Informática (SEINFO), órgãos responsáveis pelo planejamento e execução das atividades concernentes à informatização do Tribunal. O SEINFO é composto por sete setores e 23 funcionários, sendo os setores:

- ◊ Setor de administração de bancos de dados;
- ◊ Setor de administração de sistemas operacionais;
- ◊ Setor de desenvolvimento de sistemas administrativos;
- ◊ Setor de desenvolvimento de sistemas judiciários;

- ◊ Setor de internet e intranet;
- ◊ Setor de suporte ao usuário; e
- ◊ Setor de treinamento e informação.

Neste contexto, o montante de equipamentos do Tribunal é de trinta servidores e quinhentos PCs.

5.2 Definição do modelo

No item 3.10 deste trabalho, tem-se como principais tópicos para a definição de um *data warehouse*: 1) determinar os dados que popularão o *data warehouse* e quais as fontes de dados que serão necessárias; 2) armazenar e gerenciar estes dados, garantindo confiabilidade e acesso adequados; e 3) a disponibilização destes dados para o suporte à decisão. Aqui, será definido apenas o modelo do *data warehouse*. Embora sejam definidas as ferramentas e fontes de dados, não serão apresentados resultados de implantação nem detalhes quanto a equipamentos, rotinas de extração e recuperação de dados. Será um modelo lógico capaz de servir de guia para uma implantação posterior. Para a definição e apresentação das definições compreendidas nos dois primeiros passos, serão seguidos os nove pontos apresentados por Kimball (1998) e referenciados no item 3.10 deste trabalho.

5.2.1 Definição dos processos

A descrição do Tribunal, sua estrutura e competência, já foram apresentados neste trabalho. Neste contexto, o SEINFO é responsável pelo planejamento e coordenação do processo de informatização do TRT, pesquisa e análise de novas tecnologias e serviços, além de suporte às atividades de informática da 12ª Região. Compete também ao Serviço de Informática, o levantamento de necessidades de recursos dentre as diversas áreas que compõe o Tribunal, designando quantitativos e especificações técnicas acerca dos equipamentos e *softwares* à serem adquiridos.

A definição dos processos executados pela área de informática fica mais compreensível a partir da descrição do ambiente de informática do TRT, abrangendo a rede do Tribunal, os sistemas e os equipamentos.

Para uma descrição resumida da rede de informática do Tribunal, serão utilizadas como base, as unidades do mesmo. Cada Foro (JCJ ou conjunto de JCJs com o respectivo Serviço de Distribuição) possui uma rede local com servidor e sistemas rodando de forma independente. No prédio sede do Tribunal está instalada uma rede local com servidores específicos para o atendimento às principais áreas do TRT, administrativa, judiciária, gabinetes de juizes, *internet/intranet*, e banco de dados, todos interligados. O Tribunal e todas as JCJs estão interligados formando a rede de computadores WAN⁵ da 12ª Região, possibilitando transferências de dados entre os vários órgãos do TRT e o acesso de todos à *internet* e a *intranet* do Tribunal.

A informatização do Tribunal é um processo bastante recente (em 1990 o TRT possuía apenas três microcomputadores PC-XT e terminais de acesso a provedores de sistemas externos como o SERPRO). Além de recente, este processo é baseado principalmente em microinformática, o que aliado à dispersão do quadro (existem analistas e programadores lotados fora do SEINFO para o atendimento direto a necessidades de algumas áreas), ocasionou o desenvolvimento de sistemas de informações isolados, sendo utilizada uma grande variedade de linguagens. Com a satisfação das necessidades operacionais básicas de cada setor, os administradores começam a sentir a necessidade de acesso às informações, a perceber a falta de integração e compartilhamento dos dados armazenados de forma isolada e independente. Além dos problemas de integração e disponibilidade dos dados, este processo gera dificuldades de manutenção dos sistemas em função da variedade de linguagens e ambientes a serem dominados pelos técnicos.

Atualmente, podem ser vistos como pontos principais a serem trabalhados pela área de informática, questões concernentes à:

⁵ Wide area network.

- ◊ expansão dos sistemas operacionais existentes e desenvolvimento de novos sistemas para áreas ainda não atendidas;
- ◊ padronização das ferramentas de informática utilizadas pelo Tribunal, preferencialmente adotando os padrões definidos nos Encontros de Diretores de Informática de TRTs, para favorecer o intercâmbio entre os Tribunais;
- ◊ integração entre os sistemas das diversas áreas com a padronização e compartilhamento de dados comuns, tais como os referentes a setores, funcionários, materiais, entre outros;
- ◊ expansão das formas de consulta e disponibilização dos dados, abrangendo não somente a criação de bases e formas de consulta para os gerentes, mas também para os funcionários e o público em geral.

Todos estes processos estão atualmente em andamento no Tribunal, e geram outros processos tais como: aquisição e distribuição de *hardware* e *software*; e treinamento de usuários. Para os propósitos deste trabalho, as ações da área de informática foram identificadas e agrupadas nos três tipos de processos que seguem:

- ◊ Projetos, tais como: desenvolvimento de sistemas; migração de sistemas; manutenção da rede de informática; nova versão do *site internet* ou *intranet*; preparação de material para eventos e apoio a participação em eventos; ou qualquer outro projeto de informática que demande uma alocação considerável de esforços, recursos e tempo;
- ◊ Solicitações de *hardware* ou *software*, incluindo novos sistemas, e atualização ou *upgrade* tanto de *hardware* quanto de *software*; e
- ◊ Distribuição e ou instalação de *hardware* e *software*.

Dos processos citados, para este projeto, optou-se por modelar a distribuição e instalação de *hardware* e *software*. Este processo foi selecionado e considerado especialmente interessante por não haver nenhum sistema operacional que contemple todas as informações envolvidas, caracterizando uma maior necessidade da solução e pela variedade de fontes de informação, o que torna o modelo mais significativo. Além de registrar as ações de distribuição, será mantido

um registro de *status* de distribuição de *hardware* e *software*, para permitir a consulta rápida à configuração encontrada em uma determinada época. Detalhando um pouco mais, o processo de distribuição de *hardware* e *software* pode ser decorrente de:

- ◊ solicitação de usuário ou setor;
- ◊ implantação de novo sistema ou recurso;
- ◊ execução de um projeto; e
- ◊ *upgrade* - atualização do parque de informática ou adequação à novas necessidades.

A distribuição (instalação) pode ser executada tanto pelo SEINFO quanto pelos programadores lotados nas distribuições ou por pessoal externo ao Tribunal.

Atualmente, as informações referentes a este processo estão bastante dispersas. O controle dos equipamentos de informática é feito pelo Serviço de Material e Patrimônio (SEMAP), os computadores, impressoras, e demais equipamentos são tratados como material permanente, da mesma forma que mesas, cadeiras ou condicionadores de ar. Quanto aos *softwares*, são adquiridos como serviços, sendo sua instalação e uso controlados apenas pelo SEINFO. Além dos equipamentos e sistemas adquiridos, o SEINFO tem a necessidade de controlar as necessidades e o uso de sistemas desenvolvidos internamente. Um fator que facilita bastante este processo é que, quanto aos equipamentos, devido às características do processo de informatização do Tribunal, da sua estrutura, e questões ligadas ao orçamento e às licitações, o TRT possui grandes lotes de máquinas iguais, facilitando controles, tais como o de prazos de garantia ou a necessidade de atualização, aumento de memória, etc.

Da definição dos processos de informática e da seleção do processo a modelar, foram identificados grupos de fatos a registrar:

- ◊ fatos de distribuição de *hardware* e *software*; e
- ◊ instantâneos de distribuição de *hardware* e *software*.

Os fatos de distribuição serão caracterizados pela definição do item distribuído, época e local de instalação do item, que correspondem às dimensões destes. Os instantâneos, possuem praticamente as mesmas dimensões dos fatos de distribuição, mas ao invés de retratarem uma mudança, uma ação de distribuição, definem o *status* do parque de *hardware* e *software* do Tribunal em determinada época.

Dos dados a serem armazenados sobre estes fatos, alguns serão transferidos periodicamente dos sistemas de material e patrimônio, e outros serão incluídos a partir de um sistema simples, a ser desenvolvido com o intuito de substituir as planilhas e tabelas atualmente mantidas de forma independente, pelos vários setores do SEINFO, para o controle da instalação e versões dos *softwares* sob sua responsabilidade. Este sistema não armazenará seus dados diretamente no *data mart*, terá suas próprias estruturas de armazenamento de dados, e seus dados transferidos para o *data mart*, da mesma forma que os dos demais sistemas, conforme políticas definidas no item 5.9.

5.2.2 A granularidade das tabelas de fatos

Cada registro em uma tabela de fatos é um grão. A granularidade definirá o maior nível de detalhe do fato. Por exemplo, com a definição de que a granularidade da tabela de fatos de distribuição seja mês-item-local, pode ser recuperado o total de equipamentos distribuídos a um local por semestre ou por ano, mas não por quinzena ou alguma data específica. Esta definição tem um impacto significativo no tamanho do *data mart*.

O grão corresponde à representação do processo na tabela de fatos do *data mart*. Para os fatos de distribuição, será criada na tabela de fatos, uma ocorrência para cada item distribuído. Ainda que vários itens sejam entregues e utilizados em conjunto, como por exemplo, a distribuição de um PC com Windows, Word, SAP1, serão geradas uma ocorrência para o equipamento e uma para cada um dos *softwares* instalados.

Para os instantâneos serão criadas ocorrências em períodos definidos para cada item instalado. Por exemplo, em um instantâneo de hoje, seria criada uma ocorrência para cada relação equipamento-*software*-local instalado no Tribunal, neste caso, um equipamento instalado no SEINFO, contendo Windows95, Word97 e Visual Basic 6.0, geraria três ocorrências, cada uma delas associando o equipamento à data atual, ao SEINFO e a um dos *softwares* nele instalados. Serão mantidos instantâneos mensais do ano corrente, sendo que em março do ano seguinte, será feita uma reorganização dos dados, mantendo somente os registros referentes a dezembro e março se for final de gestão da Presidência do TRT.

A definição do grão, como mencionado, determina as dimensões primárias da tabela de fatos, normalmente, será possível adicionar outras dimensões ao grão básico da tabela de fatos, sendo que essas dimensões adicionais devem produzir um único valor para cada combinação das dimensões primárias. Se uma nova dimensão modificar a granularidade da tabela, gerando registros adicionais, a definição do grão deve ser adequada para acomodar a dimensão adicional.

5.2.3 As dimensões de cada tabela de fatos

As dimensões correspondem aos elementos que caracterizam a ocorrência de um fato, sendo estas características, os futuros critérios de seleções, restrições, condições de agrupamento, enfim, os parâmetros das consultas aos dados armazenados.

Nos fatos de distribuição são aplicadas as seguintes dimensões (figura 3):

Tempo	os dados em um <i>data warehouse</i> típico são construídos em torno da dimensão tempo, o tempo é o critério de seleção primário em uma grande percentagem de todas as atividades sobre o <i>data warehouse</i> . Consultas comuns buscam a atividade anual, o movimento de um mês ou semana
-------	--

Item distribuído (hardware/software)	caracterização do item, indicando se o mesmo é <i>hardware</i> ou <i>software</i> , marca, fabricante, tombo, nº de licença, enfim, esta dimensão pretente definir o item com todos os atributos necessários para sua identificação e recuperação
Local de instalação	identifica todas as unidades do TRT, e os Juizes (recebedores dos itens distribuídos). Corresponde a identificação do setor ou responsável existente no sistema de controle de materiais
Instalador	corresponde à identificação do executor da instalação, que pode ser um setor do Tribunal ou uma empresa externa contratada para tal
Projeto	identifica os projetos executado pelo SEINFO, no sentido exposto no item 5.1 (definição dos processos)

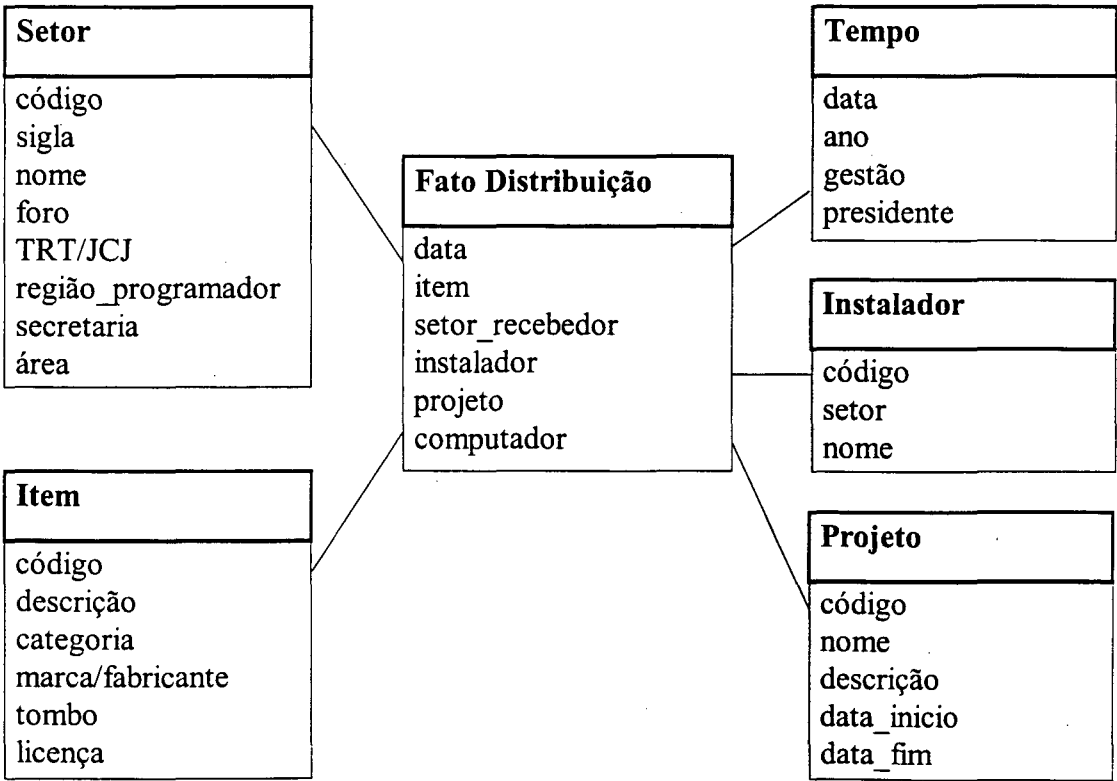


Figura 3 - Modelo de processos de distribuição

Nos instantâneos de distribuição são aplicadas as seguintes dimensões (figura 4):

Tempo	dimensão análoga à da tabela de fatos de distribuição, caracteriza a época retratada em cada instantâneo armazenado
Equipamento	identifica o equipamento (<i>hardware</i>), contendo características técnicas e informações tais como tombo, marca e garantia
Software	define os sistemas existentes no Tribunal, sejam eles adquiridos de terceiros ou desenvolvidos internamente
Local	corresponde a dimensão setor da tabela de fatos de distribuição, identificando, neste caso, o local de instalação do equipamento

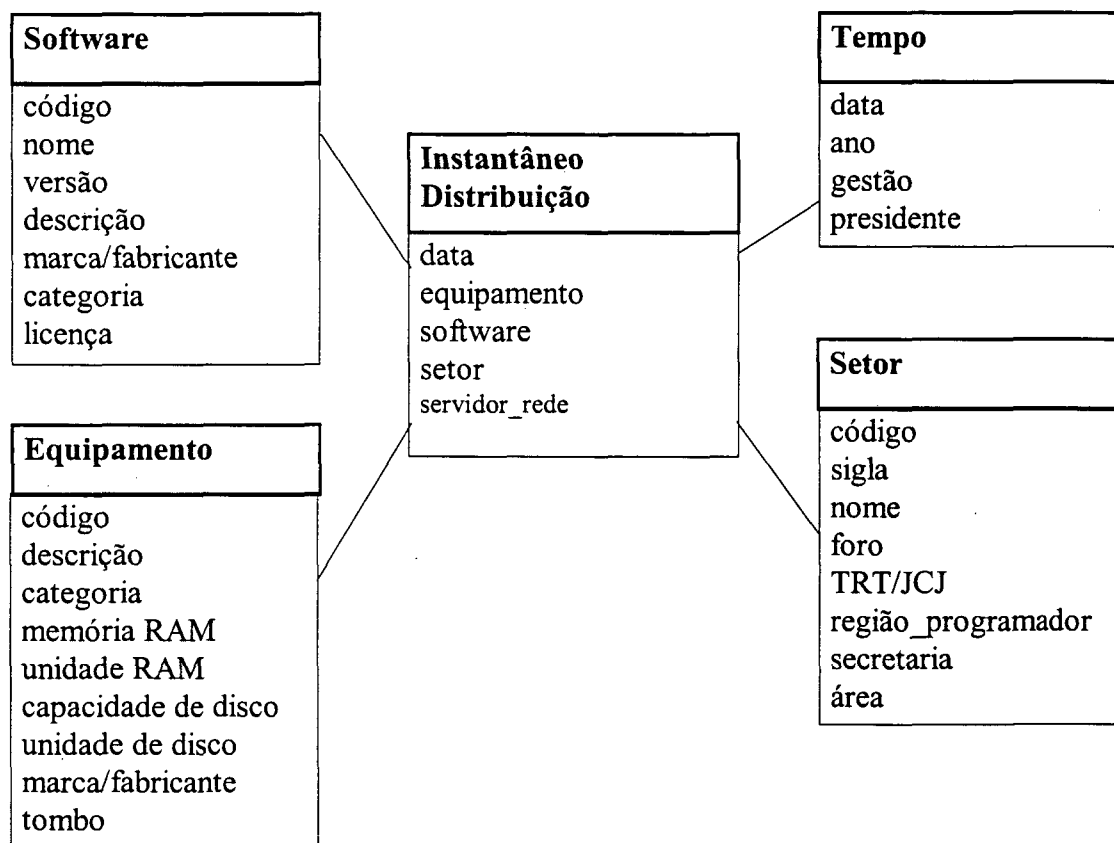


Figura 4 - Modelo de instantâneos de distribuição

5.2.4 Definição dos fatos armazenados

Os fatos correspondem às medições do negócio armazenadas no data mart, sendo um fato caracterizado pela interseção das dimensões definidas para a tabela em questão. No caso da tabela de distribuição, cada fato identificará um item distribuído em uma data, para um setor, instalado por um responsável, como parte da execução de um projeto (item x tempo x recebedor x instalador x projeto). Por sua vez, os instantâneos de distribuição definem a existência em uma determinada data, de um equipamento, em um local, com um software instalado (tempo x equipamento x local x *software*).

5.2.5 Os atributos de cada dimensão

O projeto das tabelas de dimensão requer cuidado especial, pois estas tabelas é que proverão o *data warehouse* ou não de flexibilidade nas consultas. É nas tabelas de dimensão que o usuário define restrições e seleciona os cabeçalhos das linhas dos relatórios gerados. A pesquisa, frequentemente, envolve a definição de restrições para um ou mais atributos de dimensão e a observação dos diferentes valores de outro atributo na presença destas restrições.

A seguir, serão descritos, por dimensão aplicada aos fatos distribuição, os atributos que serão armazenados.

Dimensão Tempo

data	dia, mês e ano
ano	ano, com quatro dígitos
gestão	biênio correspondente à gestão do presidente do TRT, ex: 1998/1999
presidente	abreviatura do nome do Juiz Presidente no período (existe uma tabela de abreviaturas de nomes dos Juizes utilizada nos relatórios dos sistemas)

Dimensão Item

código	código numérico seqüencial usado internamente pelo sistema para identificação do item e relacionamento com a tabela de fatos
descrição	descrição textual do item, extraída do sistema de materiais
categoria	agrupamento dos itens conforme função, produzida durante o procedimento de carga do <i>data mart</i> , divide os itens em categorias tais como: <i>hardware</i> servidor, <i>hardware desktop</i> , <i>hardware</i> comunicação (hub, roteador), <i>software</i> 3º servidor, <i>software</i> 3º cliente, <i>software</i> TRT
marca/fabricante	identificação da marca ou fabricante do item, extraída do sistema de materiais
tombo	número de tombo do equipamento no cadastro de materiais do Tribunal, também extraído do sistema de materiais (não se aplica a <i>software</i>)
licença	número da licença de uso do software (dado pelo fabricante)

Dimensão Setor

código	código numérico seqüencial para uso interno do sistema
sigla	sigla de identificação do setor conforme padrão adotado pelo TRT
nome	nome do setor
foro	atributo exclusivo de JCJ e UJA, corresponde ao código da distribuição ou JCJ à qual a JCJ ou UJA está vinculada
TRT/JCJ	indicador de localização do setor, se na sede do Tribunal ou não
região programador	atributo somente para JCJs e UJAs, correspondendo ao código da distribuição onde está lotado o programador responsável pelo atendimento ao setor

secretaria	código da secretaria ao qual o setor está vinculado na estrutura do TRT
área	identifica se o setor pertence á área administrativa ou judiciária do TRT

Dimensão Instalador

código	código numérico seqüencial para uso interno do sistema
setor	código do setor responsável pela instalação
nome	identificação do responsável pela instalação, correspondendo ao nome do setor se o instalador pertencer ao Tribunal ou nome da empresa instaladora

Dimensão Projeto

código	código numérico seqüencial para uso interno do sistema
nome	identificação do projeto
descrição	texto descritivo breve do projeto
data inicio	data de inicio da execução do projeto
data fim	data conclusão da execução do projeto

Dimensões aplicadas aos instantâneos de distribuição:

Dimensão Tempo mesma dimensão dos fatos de distribuição

Dimensão Local corresponde à dimensão setor, dos fatos de distribuição

Dimensão *Software*

código	código numérico seqüencial para uso interno do sistema
descrição	descrição textual do item, extraída do sistema de materiais
marca/fabricante	identificação da marca ou fabricante do item, extraída do sistema de materiais
licença	número da licença de uso do <i>software</i>

categoria	agrupamento dos itens conforme função, produzida durante o procedimento de carga do <i>data mart</i>
-----------	--

Dimensão Equipamento

código	código numérico seqüencial para uso interno do sistema
descrição	descrição textual do item, extraída do sistema de materiais
categoria	agrupamento dos itens conforme função, produzida durante o procedimento de carga do <i>data mart</i>
memória RAM	quantidade de memória RAM do equipamento
unidade RAM	indica a unidade em que está definido o valor de memória RAM (Mb)
capacidade de disco	capacidade total, se o equipamento possuir mais de um disco instalado
unidade disco	indica a unidade em que está definido o valor da capacidade de disco (Gb)
tombo	número de tombo do equipamento no cadastro de materiais do Tribunal, também extraído do sistema de materiais
marca/fabricante	identificação da marca ou fabricante do item, extraída do sistema de materiais

5.2.6 Forma de tratamento das dimensões de modificação lenta

Como apresentado no item 3.10, algumas dimensões variam com o tempo, e da forma de tratamento selecionada para estas dimensões, vai depender a forma de consulta e rastreamento dos históricos destas variações. Neste projeto, as dimensões de modificação lenta, correspondem às dimensões item e setor, que serão armazenadas como dimensões de modificação lenta tipo 2, ou seja, sempre que houver qualquer modificação na estrutura do Tribunal, como a extinção ou modificação de um setor, será criado um novo registro na

tabela de dimensão, para manutenção de históricos consistentes, o mesmo ocorrendo quando um item sofrer *upgrade*.

5.2.7 Definição de agregados

Agregados, são estruturas criadas para manter dados pré-calculados e pré-armazenados no *data warehouse*. A manutenção de agregados permite o acesso direto a dados anteriormente calculados, o que aumenta a performance das consultas, evitando a execução de operações de cálculo de totais, médias, valores máximos e mínimos utilizados com frequência nas consultas. Inicialmente, não serão criados agregados, pois o volume de dados do *data mart* será compatível com a execução de todas as operações diretamente sobre as tabelas.

5.2.8 Amplitude de tempo do histórico armazenado no data mart

Nos sistemas operacionais, os dados são arquivados após tornarem-se inativos por algum tempo. A razão principal para este arquivamento tem sido a performance destes sistemas. Para *data warehouses*, o projeto pode iniciar sem um plano de arquivamento de dados, pois o custo de manutenção dos dados após sua carga no *data warehouse* é mínimo. (Gupta, 1997) Como estimativa, está previsto o armazenamento de 6 (seis) anos de histórico, pois dificilmente um equipamento ou *software* mantém-se em uso por período superior a este, e as gestões do TRT correspondem a 2 (dois) anos. Estando sempre disponível, desta forma, o histórico das ações das duas últimas administrações, para as análises que se fizerem necessárias.

5.2.9 Intervalos de extração de dados e carga do data mart

A situação ideal é que os dados sejam mantidos o mais atualizados possível. Neste caso, optou-se por carregar os dados diariamente.

Esta política é viável, pois o volume de dados é pequeno e concentrado nos servidores da sede do Tribunal, de modo que o processo de extração e transferência não tende a gerar impacto no funcionamento dos sistemas ou da rede do TRT.

5.2.10 A disponibilização dos dados

Tipicamente, os dados armazenados em um *data warehouse* são disponibilizados através de três tipos de *interfaces*: a primeira voltada aos executivos, oferecendo relatórios predefinidos, com possibilidades de *drill down* e alertas destacando resultados não usuais; a segunda, destinada ao analista de negócios, ao qual são oferecidos recursos para gerar e modificar relatórios predefinidos e parametrizados; e a terceira é a interface do desenvolvedor, que contempla ferramentas para a construção dos relatórios.

Para a definição das ferramentas de consulta a serem utilizadas deve-se observar que as atividades fundamentais dos usuários de *data warehouses* são: comparação, apresentação, e questionamento do porquê dos dados resultantes. Além disto, a ferramenta deve ser de fácil utilização, neste caso, a facilidade de uso é mais importante que a disponibilidade de recursos, sendo que as consultas mais comuns ao *data warehouse* devem ser disponibilizadas em relatórios predefinidos e parametrizados.

Neste projeto, para acesso aos dados do *data mart* serão disponibilizadas via *web browser*, inicialmente consultas predefinidas. Aos poucos, com a avaliação do uso e da performance, e a partir do *feedback* dos usuários serão desenvolvidos e disponibilizados relatórios mais flexíveis, sendo posteriormente desenvolvida uma ferramenta para a execução de *queries ad hoc*. Todas as consultas serão desenvolvidas utilizando a mesma tecnologia atualmente empregada nas consultas aos bancos de dados do Tribunal disponíveis na *internet*, desenvolvidas em ASP e Visual Basic.

5.3 Limitações do projeto

Em função dos padrões de *software* e *hardware* adotados pelo Tribunal, este projeto será implementado utilizando servidor de banco de dados com Novell Netware 4.11 e Oracle7 Server release 7.3.4, sendo as estações clientes PCs com Windows 95 e as consultas disponibilizadas utilizando tecnologia *intranet*, sendo acessíveis através de *web browser* (Internet Explorer) .

5.4 Tendências da área no Tribunal

Dentro da área de abrangência deste projeto, tem-se no Tribunal demandas conhecidas por *data marts* para suprimento de informações referentes a material e patrimônio; movimento processual e outras atividades das JCJs; e movimento processual no Tribunal. Na medida em que estes forem sendo implantados e os resultados começarem a ser usufruídos, certamente surgirão novas demandas, principalmente na área administrativa do Tribunal.

A favorecer este processo tem-se a migração dos sistemas existentes para a plataforma padrão adotada (Oracle) o que facilitará em muito a transferência de dados e o tratamento dos mesmos.

Também contribui para este processo a ênfase da administração na integração das diversas áreas do Tribunal e JCJs e na disponibilização de recursos através da *internet* e *intranet*, bem como a redução da impressão de relatórios e tramitação de papéis que vem sendo substituídos pela utilização de *e-mail* ou outras formas de comunicação e transferência de dados para outros órgãos, como a Imprensa Oficial.

Tem sido observado também o reconhecimento e apoio ao trabalho fora do ambiente do Tribunal, não só dos Juizes Substitutos, que são lotados por região, mas de grande parte dos funcionários, através da disponibilização de correio eletrônico, acesso doméstico a *internet/intranet*, aquisição de *notebooks* e crescente disponibilização de recursos na *intranet*.

Finalmente, a situação de redução de verbas e o grau de informatização alcançado pelo Tribunal têm gerado um maior interesse na utilização mais racional dos recursos (financeiros e de informática), o que pode ser observado pelos estudos de O&M em andamento com objetivo de propiciar uma adequação melhor dos procedimentos aos recursos disponíveis e alocação otimizada dos recursos e sistemas.

CONSIDERAÇÕES FINAIS

Há cem anos atras, carvão e óleo tornaram-se a matéria-prima essencial para o abastecimento das fábricas da era industrial. Atualmente, a matéria-prima que alimenta as organizações e proporciona sua posição competitiva no mercado global é a informação. Sem informação, e acesso apropriado a esta, não se pode competir mais que uma fábrica de aço sem carvão pode produzir aço. A questão atual é: as organizações dispõem de informações em abundância. No entanto, poucas fazem uso de dispositivos eficazes e eficientes para localizar a informação necessária e ideal à tomada de decisão, e nem mesmo de ferramentas adequadas para transforma-la no combustível necessário para competir e sobreviver. (Gill, 1999)

Neste trabalho, foram estudadas tecnologias de *data warehousing*, que estão se tornando ferramentas importantes para a disponibilização destas informações. Estas tecnologias, que inicialmente podem ser consideradas de desenvolvimento caro, de implementação demorada e adequadas ou acessíveis somente a grandes corporações, podem ser implantadas e começar a produzir resultados rapidamente em empresas de qualquer porte. O uso de tecnologias de informação deste tipo em uma empresa, pode ser iniciado a partir da seleção e modelagem de um processo (geralmente é selecionado o processo com dados mais importantes para análise do negócio, para que os resultados encorajem a continuidade do projeto), o que diz respeito ao denominado *data mart*. A este *data mart* inicial, aos poucos, devem ser adicionados novos dados e processos, podendo-se, ao final, produzir tanto um *data warehouse* centralizado, quanto vários *data marts* departamentais, que possam ser consultados de forma independente ou integrada.

Neste contexto, para o desenvolvimento deste trabalho, foi escolhido e modelado o processo de distribuição de equipamentos e *software*, dentre os processos executados pela área de informática do TRT - 12ª Região. Com este projeto, procurou-se empregar os conhecimentos adquiridos, prestando-se este trabalho não somente ao armazenamento e disponibilização das informações sobre os processos modelados, mas também propiciando o domínio das tecnologias envolvidas para posterior emprego em outros projetos que, com certeza, virão, como pode ser percebido na mensagem do Presidente na página *internet* do Tribunal.

Bem-vindo à Home Page do Tribunal. Temos o prazer de recebê-lo, almejando que a informação pesquisada atenda à sua expectativa.

A internet está para a informação, o conhecimento e a cultura como foi o primeiro livro impresso graficamente. Aquele que não tiver a audácia de navegar pela informação precisa e imediata do sistema, sofrerá as consequências modernas do desconhecimento atualizado da realidade global.

Antonio Carlos Facioli Chedid - Juiz Presidente

O Tribunal tem procurado, em suas administrações, utilizar-se da melhor forma possível, dos recursos tecnológicos que vão sendo disponibilizados para uma prestação cada vez melhor de seus serviços à comunidade, mas ainda há muito a ser feito.

REFERÊNCIAS BIBLIOGRÁFICAS

- CHIAVENATO, Idalberto. **Teoria geral da administração**. 3 ed. São Paulo: McGraw-Hill do Brasil, 1987, vol. I.
- CONTADOR, José Celso. **Recomendações sobre o processo de Planejamento Estratégico**. RAE - Revista de Administração de Empresas, São Paulo, v. 35, n. 3, p. 39-48, Mai. Jun. 1995.
- DEGEN, Ronald. **O empreendedor** - fundamentos da iniciativa empresarial. 4 ed. São Paulo: McGraw-Hill, 1989.
- DRUCKER, Peter F. **A administração, essa desconhecida**. Exame, São Paulo, p. 60-64, 19/outubro/1988.
- FILHO, Carlos Roberto de A. R., DANTAS, Christine Ribeiro. **Tomada de Decisão Sustentada por Data warehouse**. <http://www.nce.ufrj.br/~mluiza/dataware/pages/home.htm>, 19/05/1999.
- FURLAN, José Davi, IVO, Ivonildo da Motta, AMARAL, Francisco Piedade. **Sistemas de Informação Executiva = EIS - Executive information systems**: como integrar os executivos ao sistema informacional das empresas. São Paulo: Makron books, 1994.
- FURLAN, José Davi. **Como elaborar e implementar o planejamento estratégico de sistemas de informação**. São Paulo: Makron, McGraw-Hill, 1991.
- GILL, Philip J. **Building Intelligent Enterprises**. Oracle magazine. July/August 1999.
- GREGO, Mauricio. **Contas Certas**. Infoexame. São Paulo, ano 13, nº 152 p. 120, nov 98.
- GUPTA, Vivek R. **An Introduction to Data Warehousing**. System Services Corporation, Chigago, Illinois, 1997. www.system-services.com.
- GUROVITZ, Helio. **Delete-se**. Fato: bilhões de dólares são torrados em investimentos inúteis em tecnologia. Questão: dá para se defender? Exame. São Paulo, ano 30 n. 12 p. 86-95, 4/junho/97.
- HOLTZ, Herman. **Databased Marketing**. São Paulo: Makron Books, 1994.
- JONES, Greg. **Data Marts: What's in a name?** Oracle magazine. January/February 1998.
- KIMBALL, Ralph. **Data warehouse toolkit**. São Paulo: Makron Books, 1998.

- KOTLER, Philip. **Administração de marketing: análise, planejamento, implementação e controle**; tradução Ailton Bonfim Brandão. Revista técnica José Antonio Gullo. 2 ed. São Paulo: Atlas, 1991.
- MACHADO, Carlos. **Como dar o tiro certo na hora de decidir**. Informática Exame. São Paulo, ano 11 n. 120 p. 48-54, março/1996.
- MALONE, Michael S. **A história dos modismos**. Informática Exame. São Paulo, ano 12 n. 134 p. 106, maio/1997.
- MILEY, Michael. **Bringing the World to your Warehouse**. Oracle magazine. September/October 1997.
- MILITELLO, Katia. **Pare de jogar dinheiro no lixo**. Informática Exame. São Paulo, ano 12 n. 135 p. 39-48, junho/1997.
- OLIVEIRA, Djalma de Pinho Rebouças de. **Sistemas de informações gerenciais: estratégicas, táticas, operacionais**. 2 ed. São Paulo: Atlas, 1993.
- PENTEADO, Sônia. **Produtividade movida a bits**. Informática Exame. São Paulo, ano 11 n. 123 p. 52-58, junho/1996.
- STONER, J. A. F. **Administração**. Rio de Janeiro: Prentice-Hall do Brasil, 1985.
- TECNOLOGIAS *Data Warehouse* e OLAP. www.jacques.ic.cti.br/ic/pqps/atps/dw.htm 19/05/1999.
- UNICAMP. Informativo Técnico nº 12. www.revista.unicamp.br/revista/infotec/informacao/inf12.htm, 1996a.
- _____. Informativo Técnico nº 18. www.revista.unicamp.br/revista/infotec/informacao/inf18.htm, 1996b.